

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 14, 2003, 13:03:37 ; Search time 94 Seconds  
(without alignments)  
494.143 Million cell updates/sec

Title: US-09-876-997-399

Perfect score: 968

Sequence: 1 MWLYRNPVBAEYFPTKPMF.....RTCDYKHHWODLLKCTNTAK 180

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 23:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp Vertebrate:\*
- 14: sp Unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	926	95.7	223	4 Q8NEB5	Q8neb5 homo sapien
2	916	94.6	175	4 Q9BY45	Q9by45 homo sapien
3	440	45.5	363	5 Q8IGV3	Q8igv3 drosophila
4	440	45.5	412	5 Q9VND5	Q9vnds drosophila
5	382	39.5	314	10 Q9M882	Q9m882 arabidopsis
6	382	39.5	364	10 Q8LFD1	Q8lfd1 arabidopsis
7	379.5	39.2	317	10 Q8R297	Q8r297 oryza sativ
8	379.5	39.2	328	10 Q94E07	Q94e07 oryza sativ
9	372	38.4	290	10 Q9X160	Q9x160 arabidopsis
10	364	37.6	374	10 Q9LLQ7	Q9llq7 vigna ungui
11	362	37.4	362	10 Q9AWT8	Q9awt8 oryza sativ
12	346	35.7	322	10 Q9FVL1	Q9fvl1 vigna ungui
13	337	34.8	302	10 Q8H0G3	Q8hog3 arabidopsis
14	337	34.8	327	10 Q8LDP8	Q8ldp8 arabidopsis
15	337	34.8	327	10 Q9ZU49	Q9zu49 arabidopsis
16	336.5	34.8	308	10 Q8LAS9	Q8las9 arabidopsis

17	336	34.7	307	10 Q9LJQ8	Q9ljq8 arabidopsis
18	335	34.6	327	10 Q945N3	Q945n3 arabidopsis
19	247	25.5	289	3 Q05521	Q05521 saccharomyc
20	245	25.3	281	5 Q8IPQ4	Q8ipq4 drosophila
21	242	25.0	279	3 Q9UUA6	Q9uua6 schizosacch
22	205.5	21.2	328	10 Q8RZY0	Q8ryz0 oryza sativ
23	200.5	20.7	274	3 Q04396	Q04396 saccharomyc
24	196	20.2	335	5 Q8T2M2	Q8t2m2 dictyostell
25	194.5	20.1	305	5 Q9VNT8	Q9vnt8 drosophila
26	188	19.4	340	5 Q8IA51	Q8ia51 caenorhabdi
27	188	19.4	346	5 Q9TXU1	Q9txu1 caenorhabdi
28	188	19.4	385	5 Q8IA52	Q8ia52 caenorhabdi
29	184.5	19.1	282	11 Q08564	Q08564 rattus norv
30	184.5	19.1	283	11 Q8K594	Q8k594 rattus norv
31	183.5	19.0	283	11 Q61469	Q61469 mus musculu
32	183.5	19.0	284	11 Q8BPB8	Q8bpb8 mus musculu
33	183	18.9	312	11 Q99JY8	Q99jy8 mus musculu
34	179.5	18.5	350	5 Q9NGZ2	Q9ngz2 drosophila
35	179.5	18.5	350	5 Q9V577	Q9v577 drosophila
36	179.5	18.5	350	5 Q8IGU7	Q8igu7 drosophila
37	179	18.5	312	11 P97544	P97544 rattus norv
38	178.5	18.4	288	4 Q43688	Q43688 homo sapien
39	174.5	18.0	286	11 Q88957	Q88957 cavia porce
40	173.5	17.9	285	11 Q88956	Q88956 cavia porce
41	173.5	17.9	311	4 Q96GW0	Q96gw0 homo sapien
42	173.5	17.9	311	4 Q14495	Q14495 homo sapien
43	172.5	17.8	284	4 Q14494	Q14494 homo sapien
44	172.5	17.8	285	4 Q60457	Q60457 homo sapien
45	172.5	17.8	289	4 Q60463	Q60463 homo sapien

ALIGNMENTS

RESULT 1

Q8NEB5 ID Q8NEB5 PRELIMINARY; PRT; 223 AA.  
AC Q8NEB5;  
DT 01-OCT-2002 (Tremblrel. 22, Created)  
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)  
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
DE Similar to HTPAP protein.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Testis;  
RA Strausberg R.;  
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC033025; AAH3025.1; -;  
DR InterPro; IPR000326; PA\_PTPase.  
DR Pfam; PF01569; PAP2; 1.  
DR SMART; SM00014; acidPPc; 1.  
SQ SEQUENCE 223 AA; 25159 MW; 72B01C9A0DBBA1BA CRC64;

Query Match 95.7%; Score 926; DB 4; Length 223;  
Best Local Similarity 99.4%; Pred. No. 2.2e-94;  
Matches 172; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MWLYRNPVBAEYFPTKPMFVIAFLSPLSLIFLAKFLKKADTRDSRQACLAASLALANG	60
Db	1	MWLYRNPVBAEYFPTKPMFVIAFLSPLSLIFLAKFLKKADTRDSRQACLAASLALANG	60
Qy	61	VFTNTIKLIIVGRPRDPFYRCFPDGLAHSDIMCTGDXVNEGRKSPFSGHSSAFAGLA	120
Db	61	VFTNTIKLIIVGRPRDPFYRCFPDGLAHSDIMCTGDXVNEGRKSPFSGHSSAFAGLA	120
Qy	121	PASFYLAGLKHCFPTQGRGKSWRCFAFLSPLLFAAVIALSRTCDYKHHWODLL	173
Db	121	PASFYLAGLKHCFPTQGRGKSWRCFAFLSPLLFAAVIALSRTCDYKHHWODLV	173

```
RESULT 2
Q9BY45 Q9BY45 PRELIMINARY; PRT; 175 AA.
ID Q9BY45 Q9BY45 PRELIMINARY; PRT; 175 AA.
AC Q9BY45;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE HTPAP.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Li Y., Wu T., Xu S., Ren S., Chen Z., Han Z.;
RT "A novel gene expressed in human liver non-tumor tissue.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF212238; AAK14924.1; -.
DR InterPro: IPR000326; PA_PTPase.
DR Pfam: PF01569; PAP2; 1.
DR SMART: SM00014; acdPPc; 1.
SQ SEQUENCE 175 AA; 19766 MW; 0133956D40539F83 CRC64;

Query Match 94.6%; Score 916; DB 4; Length 175;
Best Local Similarity 98.38; Pred No. 2.2e-93;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MWLYRNPYVEAEYPTKPFVIAFLPSLIFLAKFLKADTRDSROACLAAS 60
DB 1 MWLYRNPYVEAEYPTKPFVIAFLPSLIFLAKFLKADTRDSROACLAAS 60
QY 61 VFTNTIKLVGRPRDPDFYRCFPDGLAHSLMCTGDKDVNVEGRKSPSGHSSFAFAGLA 120
DB 61 VFTNTIKLVGRPRDPDFYRCFPDGLAHSLMCTGDKDVNVEGRKSPSGHSSFAFAGLA 120
QY 121 FASFLAGLKLHCFTPOGRGKSWRCFAFLSPLLPFAAVIALSRTCDYKHWWQDLK 174
DB 121 FASFLAGLKLHCFTPOGRGKSWRCFAFLSPLLPFAAVIALSRTCDYKHWWQDPFK 174

RESULT 3
Q8IGV3 Q8IGV3 PRELIMINARY; PRT; 363 AA.
ID Q8IGV3 Q8IGV3 PRELIMINARY; PRT; 363 AA.
AC Q8IGV3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE R23632p.
GN CG12746.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Y;
RA Chapote M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BT001579; AAN71334.1; -.
SQ SEQUENCE 363 AA; 41230 MW; D46A8483FD4F8E72 CRC64;

Query Match 45.5%; Score 440; DB 5; Length 363;
Best Local Similarity 44.58; Pred No. 2.8e-40;
Matches 85; Conservative 29; Mismatches 49; Indels 28; Gaps 4;

QY 1 MWLYRNPYVEAEYPTKPFVIAFLPSLIFLAKFLKADTRDSROACLAAS 53
DB 103 LWLYKN-----PRPDIVRGELLFWVIVAPP-LVTIAFYWYTRDRDFFRAASWAWT 153
QY 54 LALALAGVFTNTIKLVGRPRDPDFYRCFPDGL-----AHSDLMCTGDKDVN 101
DB 154 LALCMGIGPISVLKIVGRPRDPDFYRCFPDGLVNLNNTSGVVDTSILDNCTGLPGDIN 213
QY 102 EGRKSPSGHSSFAFAGLAFASFLYLAGLKLHCFTPOGRGKSWRCFAFLSPLLPFAAVIALSR 161
DB 214 EGRKSPSGHSSFAFAGLAFASFLYLAGLKLHCFTPOGRGKSWRCFAFLSPLLPFAAVIALSR 273
QY 162 TCDYKHWWQDL 172
DB 274 TCDYHHHWWQDV 284

RESULT 4
Q9VND5 Q9VND5 PRELIMINARY; PRT; 412 AA.
ID Q9VND5 Q9VND5 PRELIMINARY; PRT; 412 AA.
AC Q9VND5; Q9VND5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CG12746 protein.
GN CG12746.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).
RC STRAIN=BERKELEY; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vardell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler K., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacieb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RL "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195 (2000).
```

CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS, A LONG FORM (SHOWN HERE) AND A  
CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.  
DR EMBL; AE003602; AAF52007.1; -  
DR EMBL; AE003602; AAF52008.1; -  
DR FlyBase; FBgn037341; CG12746.  
DR InterPro; IPR000326; PA\_PTPase.  
DR Pfam; PF01569; PAP2; 1\_  
DR SMART; SM00014; acidPPc; 1\_  
KW Alternative splicing; Hypothetical protein.  
FT VARSPPLIC 1 49 MISSING (IN SHORT ISOFORM).  
SQ SEQUENCE 412 AA; 46736 MW; 0966C4340A444C1 CRC64;  
  
Query Match 45.5%; Score 440; DB 5; Length 412;  
Best Local Similarity 44.5%; Pred. No. 3.2e-40;  
Matches 85; Conservative 29; Mismatches 49; Indels 28; Gaps 4;  
  
QY 1 MWLYRNPVVEAEYPTKP-----MFVIAFLPSLFIKFLKADTRDSRQACLAAS 53  
DB LLYLKN-----PRPDIVRGCELLFWIVAPP-LVTIAFYWYTRDRDFRAASWAWT 202  
  
QY 54 LALALNGVFTNTIKLVGRPRDPFFRCPPDGL-----AHSDLMTGDKDQVNV 101  
DB LLYLKN-----PRPDIVRGCELLFWIVAPP-LVTIAFYWYTRDRDFRAASWAWT 202  
  
QY 203 LALCMNGIPTSVLKITVGRPRDPFYRCFDPGVNVLNTSNGVDTSLDFNCTGLPGDIN 262  
DB LLYLKN-----PRPDIVRGCELLFWIVAPP-LVTIAFYWYTRDRDFRAASWAWT 202  
  
QY 102 EGRKSPFSGHSSAFAGLAFASFLACKLHCFTQGRGKSWRCFAFLSPILLFAVIALSR 161  
DB EGRKSPFSGHSSAFAGLAFASFLACKLHCFTQGRGKSWRCFAFLSPILLFAVIALSR 161  
  
QY 263 EGRKSPFSGHSSAFAGLAFASFLACKLHCFTQGRGKSWRCFAFLSPILLFAVIALSR 322  
DB EGRKSPFSGHSSAFAGLAFASFLACKLHCFTQGRGKSWRCFAFLSPILLFAVIALSR 322  
  
QY 162 TCDYKHHWQDL 172  
DB TCDYKHHWQDV 333  
  
RESULT 5  
Q9M882 ID Q9M882 PRELIMINARY; PRT; 314 AA.  
AC Q9M882;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Putative phosphatidate phosphohydrolase.  
GN F16B3.23  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,  
RA Ronning C.M., Koo H., Fujii C.Y., Uterback T.R., Barnstead M.E.,  
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;  
RT "Arabidopsis thaliana chromosome III BAC F16B3 genomic sequence."  
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC021640; AAF32467.1; -  
DR InterPro; IPR000326; PA\_PTPase.  
DR Pfam; PF01569; PAP2; 1\_  
DR SMART; SM00014; acidPPc; 1\_  
KW Hydrolase.  
SQ SEQUENCE 314 AA; 35184 MW; 5F1E88546058C497 CRC64;  
  
Query Match 39.5%; Score 382; DB 10; Length 314;  
Best Local Similarity 47.0%; Pred. No. 6.3e-34;  
Matches 77; Conservative 28; Mismatches 49; Indels 10; Gaps 4;  
  
QY 16 TKPMF---VIAFLSPSLIFLAKFLKADTRDSRQACLAASLALANGVFTNTIKLVGR 72  
DB TVPIWSPVYVAMLLPL-VIFIFIFRRRDVYDLHVAVLGLLSVLTAVLTDAIKNAVGR 123  
  
QY 73 PRDFFVRCPPDGLA----HSDLMCTGDKDVNNGRKSFPSSGHSSAFAGLAFASFLAG 128  
DB PRDFFVRCPPDGLA----HSDLMCTGDKDVNNGRKSFPSSGHSSAFAGLAFASFLAG 128  
  
QY 124 PRDFFVRCPPDGLA----HSDLMCTGDKDVNNGRKSFPSSGHSSAFAGLAFASFLAG 183  
DB PRDFFVRCPPDGLA----HSDLMCTGDKDVNNGRKSFPSSGHSSAFAGLAFASFLAG 183

QY 129 KLHCFTPQGRGKSWRCFAFLSPILLFAVIALSRTCDYKHHWQDL 172  
DB KIQAF--DGKGVHAKLCIVILPLFAALVGISRVDDYWHHWDV 225  
  
RESULT 6  
Q8LFD1 ID Q8LFD1 PRELIMINARY; PRT; 364 AA.  
AC Q8LFD1;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Putative phosphatidate phosphohydrolase.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Haas B.J., Volkovsky N., Town C.D., Troukhan M., Alexandrov N.,  
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;  
RT "Full-length messenger RNA sequences greatly improve genome  
RT annotation."  
RL Genome Biol. 0:0-0 (2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,  
RA Feldmann K.;  
RT "Full-length cDNA from Arabidopsis thaliana."  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY084915; AAM61477.1; -  
DR InterPro; IPR000326; PA\_PTPase.  
DR Pfam; PF01569; PAP2; 1\_  
DR SMART; SM00014; acidPPc; 1\_  
KW Hydrolase.  
SQ SEQUENCE 364 AA; 40775 MW; 905F149C235E4A18 CRC64;  
  
Query Match 39.5%; Score 382; DB 10; Length 364;  
Best Local Similarity 47.0%; Pred. No. 7.5e-34;  
Matches 77; Conservative 28; Mismatches 49; Indels 10; Gaps 4;  
  
QY 16 TKPMF---VIAFLSPSLIFLAKFLKADTRDSRQACLAASLALANGVFTNTIKLVGR 72  
DB TVPIWSPVYVAMLLPL-VIFIFIFRRRDVYDLHVAVLGLLSVLTAVLTDAIKNAVGR 173  
  
QY 73 PRDFFVRCPPDGLA----HSDLMCTGDKDVNNGRKSFPSSGHSSAFAGLAFASFLAG 128  
DB PRDFFVRCPPDGLA----HSDLMCTGDKDVNNGRKSFPSSGHSSAFAGLAFASFLAG 128  
  
QY 129 KLHCFTPQGRGKSWRCFAFLSPILLFAVIALSRTCDYKHHWQDL 172  
DB KIQAF--DGKGVHAKLCIVILPLFAALVGISRVDDYWHHWDV 275  
  
RESULT 7  
Q8RZ97 ID Q8RZ97 PRELIMINARY; PRT; 317 AA.  
AC Q8RZ97;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Putative phosphatidic acid phosphatase beta.  
GN OSJNB0063G05.1.  
OS Oryza sativa (Japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaceae; Oryza.  
OX NCBI\_TaxID=39947;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Nipponbare;

RA Sasaki T., Matsumoto T., Yamamoto K.;  
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, BAC  
RL clone:OSUNBB063G05";  
Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AP003760; BAB90492.1; -  
DR Gramene; Q8R297; -  
DR InterPro; IPR000326; PA\_PTPase.  
DR Pfam; PF01569; PAP2; 1.  
DR SMART; SM00014; acidPPc; 1.  
SQ SEQUENCE 317 AA; 34832 MW; 9008B4DF7FC5D849 CRC64;

Query Match 39.2%; Score 379.5; DB 10; Length 328;  
Best Local Similarity 45.0%; Pred. No. 1.3e-33;  
Matches 76; Conservative 33; Mismatches 47; Indels 13; Gaps 5;

QY 16 TKPMF---VIAFLSPLSLIFLAKFLKKADTRDSRQACLAASLALNGVFTNTIKLIVGR 72  
DB 71 TVPIWAVPIVAVIGPM-IVFTVVYFRNRVYDLHHAIVLGILFSLVLTGVLTDKDAVGR 129  
QY 73 PRDFFVRCFPDGLAHS-DMCTGDKDVNVEGRKSPSPGSHSSFAFAGLAFASFLA 127  
DB 130 PRNFFWRCFPGDIAVDNVTGVIHGDSASVIKEGHSKSPSGHTSWSFAGLGLSWYLA 189  
QY 128 GKLCFTFQGRGKSWRCFASPLSPLFAVIALSRTCDYKHHWQDLKCT 176  
DB 190 GKITVF--DRRGHVAKLCVVLAPLVAAMVAISRVDYWHWQDV--CT 234

## RESULT 8

Q94E07  
ID Q94E07 PRELIMINARY; PRT; 328 AA.  
AC Q94E07  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Putative phosphatidic acid phosphatase beta.  
GN P0003E08.23.  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaceae; Oryza.  
OX NCBI\_TaxID=4530;  
[1]  
SEQUENCE FROM N.A.  
RC STRAIN=cv. Nipponbare;  
RA Sasaki T., Matsumoto T., Yamamoto K.;  
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC  
RL clone:P0003E08";  
Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AP003222; BAB63556.1; -  
DR Gramene; Q94E07; -  
DR InterPro; IPR000326; PA\_PTPase.  
DR Pfam; PF01569; PAP2; 1.  
DR SMART; SM00014; acidPPc; 1.  
SQ SEQUENCE 328 AA; 36004 MW; 5CDD3A44ABDB9C55 CRC64;

Query Match 39.2%; Score 379.5; DB 10; Length 328;  
Best Local Similarity 45.0%; Pred. No. 1.3e-33;  
Matches 76; Conservative 33; Mismatches 47; Indels 13; Gaps 5;

QY 16 TKPMF---VIAFLSPLSLIFLAKFLKKADTRDSRQACLAASLALNGVFTNTIKLIVGR 72  
DB 71 TVPIWAVPIVAVIGPM-IVFTVVYFRNRVYDLHHAIVLGILFSLVLTGVLTDKDAVGR 129  
QY 73 PRDFFVRCFPDGLAHS-DMCTGDKDVNVEGRKSPSPGSHSSFAFAGLAFASFLA 127  
DB 130 PRNFFWRCFPGDIAVDNVTGVIHGDSASVIKEGHSKSPSGHTSWSFAGLGLSWYLA 189  
QY 128 GKLCFTFQGRGKSWRCFASPLSPLFAVIALSRTCDYKHHWQDLKCT 176  
DB 190 GKITVF--DRRGHVAKLCVVLAPLVAAMVAISRVDYWHWQDV--CT 234

RESULT 9  
Q9XI60  
ID Q9XI60 PRELIMINARY; PRT; 290 AA.  
AC Q9XI60;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE F9L1.2 protein (Putative phosphatidic acid phosphatase) (Prenyl  
DE diphosphate phosphatase).  
GN F9L1.2 OR ATPAP2;  
OS Arabidopsis thaliana (Mouse-ear cross).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
OX NCBI\_TaxID=3702;  
[1]  
SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RA Vysotskaia V.S., Schwartz J.R., Yu G., Toriumi M., Lenz C., Liu S.,  
Lee J., Liu A., Li J., Kremenetskaia I., Luros J., Gonzalez A.,  
Altafi H., Araujo R., Brooks S., Buehler E., Chao Q., Conn L.,  
Conway A.B., Dunn P., Hansen N., Huizar L., Khan S., Kim C., Palm C.,  
Rowley D., Shinn P., Walker M., Davis R.W., Ecker J.R.,  
Fiederspiel N.A., Theologis A.;  
RA "Arabidopsis thaliana chromosome 1 BAC F9L1 sequence";  
RT Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
[2]  
SEQUENCE FROM N.A.  
RC Katagiri T., Shinzaki K.;  
RA "A cDNA sequence encoding a phosphatidic acid phosphatase in  
RT Arabidopsis thaliana";  
Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
[3]  
SEQUENCE FROM N.A.  
RC Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,  
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;  
RT "Full-length messenger RNA sequences greatly improve genome  
RT annotation";  
RT Genome Biol. 0:0-0 (2002).  
[4]  
SEQUENCE FROM N.A.  
RC Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,  
RA Feldmann K.;  
RT "Full-length cDNA from Arabidopsis thaliana";  
Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
[5]  
SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RA Tokuhito K., Muramoto N., Yamada Y., Asami O., Hirai M., Obata S.,  
RA Ohto C., Muramatsu M.;  
RT "Prenyl alcohol production by overexpression of prenyl diphosphate  
RT phosphatase in Yeast Saccharomyces cerevisiae";  
Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
RL EMBL; AC007591; AAD39637.1; -  
DR EMBL; AB061407; BAB47574.1; -  
DR EMBL; AY087673; AAM65210.1; -  
DR EMBL; AB053950; BAC41335.1; -  
DR InterPro; IPR000326; PA\_PTPase.  
DR Pfam; PF01569; PAP2; 1.  
SQ SEQUENCE 290 AA; 32702 MW; BF14A9A0C23B4429 CRC64;

Query Match 38.4%; Score 372; DB 10; Length 290;  
Best Local Similarity 47.8%; Pred. No. 7.4e-33;  
Matches 75; Conservative 28; Mismatches 46; Indels 8; Gaps 4;

QY 21 VIAFLSPLSLIFLAKFLKKADTRDSRQACLAASLALNGVFTNTIKLIVGRPPDPFYR 80  
DB 73 LIAVLPFAVICVYFIRN-DVYDLHHAIVLGILFSLVLTGVLTDKDAVGRPPDPFYR 131  
QY 81 CPFDGLA--HS---DMCTGDKDVNVEGRKSPSPGSHSSFAFAGLAFASFLAGLKHCTP 135  
DB 132 CPFDGIGIFHNVTNVLCTGAKDVVKEGHSKSPSGHTSWSFAGLGLSLYLSGKIRF-- 189

QY 136 QGRGKSWRCFAFLSPPLFAAVIALSRCTCDYKHHWQDL 172  
DB 190 DQRGHVAKLCIVLPLLLVAALVGSVRVDDYWHHWDV 226

## RESULT 10

Q9LLQ7 PRELIMINARY; PRT; 374 AA.  
AC Q9LLQ7;  
DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE Phosphatidic acid phosphatase alpha (EC 3.1.3.4).  
OS Vigna unguiculata (Copeae).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids 1; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Vigna.  
OX NCBI\_TaxID=3917;  
RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN=cv. EPAGE-1; TISSUE=Leaf;  
RA Franca M.G.C., Matos A.R., d'Arcy-Lameta A., Zuily-Fodil Y.,  
RA Pham-Thi A.T.;  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF165891; AAF89579.1; -  
DR InterPro; IPR000326; PA\_PTPase.  
DR Pfam; PF01569; PAP2; 1.  
DR SMART; SM00014; acidppc; 1.  
KW Hydrolyase.  
SQ SEQUENCE 374 AA; 42336 MW; B50DADS707D0A0BC CRC64;

Query Match 37.6%; Score 364; DB 10; Length 374;  
Best Local Similarity 45.7%; Pred. No. 7.6e-32;  
Matches 75; Conservative 27; Mismatches 52; Indels 10; Gaps 4;

QY 16 TKMPF---VIAFLSPSLIFLAKFKKADTRDSRQACLAASLALANGVFTNTIKLIVGR 72  
DB 120 TVPVMSIPIYAVLLPI-VIEFLVYIRRRDYDLHFAVLGLLSLITAVITEAKNGVGR 178  
QY 73 PRPDFFVRCPPDGL---AHSIDLMTCTGDKDVNVEGRKSPFSGHSSFAFAGLAFASFYLAG 128  
DB 179 PRPDFFVRCPPDGL---AHSIDLMTCTGDKDVNVEGRKSPFSGHSSFAFAGLAFASFYLAG 238  
QY 129 KLHCFTPQGRGKSWRCFAFLSPPLFAAVIALSRCTCDYKHHWQDL 172  
DB 239 KIRAF--DRGHVAKLCIVLPLLLFASLVGISRVDYWHHWDV 280

## RESULT 11

Q9AWT8 PRELIMINARY; PRT; 362 AA.  
AC Q9AWT8;  
DT 01-JUN-2001 (Tremblrel. 17, Created)  
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)  
DE Putative phosphatidic acid phosphatase alpha.  
GN P0480E02.6.  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzeae; Oryza.  
OX NCBI\_TaxID=4530;  
RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Nipponbare;  
RA Sasaki T., Matsumoto T., Yamamoto K.;  
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC  
clone: P0480E02.6";  
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AP002913; BAB21200.1; -  
DR Gramene; Q9AWT8; -  
DR InterPro; IPR000326; PA\_PTPase.  
DR Pfam; PF01569; PAP2; 1.

DR SMART; SM00014; acidppc; 1.  
SQ SEQUENCE 362 AA; 40682 MW; 8C4F1141F8BF5176 CRC64;  
Query Match 37.4%; Score 362; DB 10; Length 362;  
Best Local Similarity 40.1%; Pred. No. 1.2e-31;  
Matches 71; Conservative 36; Mismatches 62; Indels 8; Gaps 3;  
QY 1 MMLYRNPYEAEEYPTKPMFVIAFLSPSLIFLAKFKKADTRDSRQACLAASLALANG 60  
DB 56 MMTYVSPVQKSTVPANGVPIISIVCPV-IIFLSVYIARRDVYDLHHAATLGVLSVLTA 114  
QY 61 VFTNTIKLVGRPRDPDFYRCPPDG-----LAHSDILMTCTGDKDVNVEGRKSPFSGHSSFA 115  
DB 115 VTTVVKNVAGRPDPDFWRCPPDGKQLXDQVTDGVICHGKSFLLDKGRKSPFSGHTSWS 174  
QY 116 FAGLAFASFYLAGLHCFTPQGRGKSWRCFAFLSPPLFAAVIALSRCTCDYKHHWQDL 172  
DB 175 FAGLGLSLYLSGKIKVFDROGHVA--KLCIMLPLLLIASLVGISRIDDYRHHWEDV 229

## RESULT 12

Q9FVL1 PRELIMINARY; PRT; 322 AA.  
AC Q9FVL1;  
DT 01-MAR-2001 (Tremblrel. 16, Created)  
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE Phosphatidic acid phosphatase beta (EC 3.1.3.4).  
OS Vigna unguiculata (Copeae).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids 1; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Vigna.  
OX NCBI\_TaxID=3917;  
RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN=cv. EPAGE-1; TISSUE=Leaf;  
RA Franca M.G.C., Matos A.R., d'Arcy-Lameta A., Zuily-Fodil Y.,  
RA Pham-Thi A.T.;  
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF171230; AAF89745.3; -  
DR InterPro; IPR000326; PA\_PTPase.  
DR Pfam; PF01569; PAP2; 1.  
DR SMART; SM00014; acidppc; 1.  
KW Hydrolyase.  
SQ SEQUENCE 322 AA; 36224 MW; 67B7B572DCAA1F8E CRC64;

Query Match 35.7%; Score 346; DB 10; Length 322;  
Best Local Similarity 45.9%; Pred. No. 6.4e-30;  
Matches 72; Conservative 25; Mismatches 52; Indels 8; Gaps 3;

QY 21 VIAFLSPSLIFLAKFKKADTRDSRQACLAASLALANGVFTNTIKLIVGRPRDPFYR 80  
DB 73 IVALLLPLAVFLVYVYFIRK-DVYDFHAILGLLSLITAVITDAIKDGVGRPRDPFYR 131  
QY 81 CPDGL-----AHSIDLMTCTGDKDVNVEGRKSPFSGHSSFAFAGLAFASFYLAGLHCFTP 135  
DB 132 CPDGLGVDPVTSVRCCTGDKVKEGTKVSPSGHTSWSFAGLVLSLWKLSGKIRV-- 189  
QY 136 QGRGKSWRCFAFLSPPLFAAVIALSRCTCDYKHHWQDL 172  
DB 190 DRGHVAKLCIVLPLLLFASLVGISRVDYWHHWDV 226

## RESULT 13

Q8H0G3 PRELIMINARY; PRT; 302 AA.  
AC Q8H0G3;  
DT 01-MAR-2003 (Tremblrel. 23, Created)  
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
DE Prenyl diphosphate phosphatase.  
GN ATPAP1.  
OS Arabidopsis thaliana (Mouse-ear cross).

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN=cv. Columbia;  
RC Tokuhito K., Muramatsu M., Yamada Y., Asami O., Hirai M., Obata S.,  
RA Ohto C., Muramatsu M.,  
RT "Prenyl alcohol production by overexpression of prenyl diphosphate  
RT phosphatase in yeast Saccharomyces cerevisiae";  
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB053949; BAC1334.1; -  
SQ SEQUENCE 302 AA; 33739 MW; 9E1CGD7DAFD569D6 CRC64;  
Query Match 34.8%; Score 337; DB 10; Length 302;  
Best Local Similarity 42.1%; Pred. No. 5.9e-29;  
Matches 69; Conservative 31; Mismatches 54; Indels 10; Gaps 4;  
Qy 16 TKPMF---VIAFLSPSLIFLAKFLKADTRDSRQACLAASLALNGVFTNTIKLIVGR 72  
Db 65 TVPIWSVPVAVLLPI-IVFVCFYKRTCVYDLHHSILGLLFAVLITGVITDSIKVATGR 123  
Qy 73 PRDFFVRCPPDGL-----AHSIDLCTGDKDVVNEGRKSPFSGHSSFAFAGLAFASFYLAG 128  
Db 124 PRPNFYWRCPDGGKELYDALGGVVCHGKAAEVKEGHSFPSPGHTSWSFAGLTFLSLYLSG 183  
Qy 129 KLHCFTPQGRKSWRFCAFLSPLLFAAVIALSRCTCDYKHHQDL 172  
Db 184 KIKAFNNEGHVA--KLCVLIFPPLAACLGVISRVDYVHHWQDV 225  
RESULT 14  
Q8LDP8 PRELIMINARY; PRT; 327 AA.  
AC Q8LDP8;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Putative phosphatidic acid phosphatase.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,  
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;  
RT "Full-length messenger RNA sequences greatly improve genome  
RT annotation";  
RL Genome Biol. 0:0-0(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,  
RA Feldmann K.;  
RT "Full-length cDNA from Arabidopsis thaliana";  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY085869; AAM63082.1; -  
DR InterPro; IPR000326; PA\_FTPase.  
DR Pfam; PF01569; PAP2; 1.  
DR SMART; SM00014; acidPPC; 1.  
SQ SEQUENCE 327 AA; 36646 MW; 748B2A02EC3BEA3E CRC64;  
Query Match 34.8%; Score 337; DB 10; Length 327;  
Best Local Similarity 42.1%; Pred. No. 6.4e-29;  
Matches 69; Conservative 31; Mismatches 54; Indels 10; Gaps 4;  
Qy 16 TKPMF---VIAFLSPSLIFLAKFLKADTRDSRQACLAASLALNGVFTNTIKLIVGR 72  
Db 90 TVPIWSVPVAVLLPI-IVFVCFYKRTCVYDLHHSILGLLFAVLITGVITDSIKVATGR 148  
Qy 73 PRDFFVRCPPDGL-----AHSIDLCTGDKDVVNEGRKSPFSGHSSFAFAGLAFASFYLAG 128

Db 149 PRPNFYWRCPDGGKELYDALGGVVCHGKAAEVKEGHSFPSPGHTSWSFAGLTFLSLYLSG 208  
Qy 129 KLHCFTPQGRKSWRFCAFLSPLLFAAVIALSRCTCDYKHHQDL 172  
Db 209 KIKAFNNEGHVA--KLCVLIFPPLAACLGVISRVDYVHHWQDV 250  
RESULT 15  
Q9ZU49 PRELIMINARY; PRT; 327 AA.  
AC Q9ZU49; Q94LY6;  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Putative phosphatidic acid phosphatase.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., VanAken S.E.,  
RA Barnstead M.E., Mason T.M., Bowman C.L., Ronning C.M., Benito M.-I.,  
RA Carrera A.J., Creasy T.H., Buell C.R., Town C.D., Nieman W.C.,  
RA Fraser C.W., Venter J.C.;  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Columbia;  
RA Katagiri T., Shinozaki K.;  
RT "One of the biological functions of the gene encoding phosphatidic  
RT acid phosphatase (PAP) correlates cell elongation in Arabidopsis  
RT thaliana";  
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC006200; AAD14518.2; -  
DR EMBL; AB004318; BAB47575.1; -  
DR InterPro; IPR000326; PA\_FTPase.  
DR Pfam; PF01569; PAP2; 1.  
DR SMART; SM00014; acidPPC; 1.  
SQ SEQUENCE 327 AA; 36683 MW; BAF6F345817BA245 CRC64;  
Query Match 34.8%; Score 337; DB 10; Length 327;  
Best Local Similarity 42.1%; Pred. No. 6.4e-29;  
Matches 69; Conservative 31; Mismatches 54; Indels 10; Gaps 4;  
Qy 16 TKPMF---VIAFLSPSLIFLAKFLKADTRDSRQACLAASLALNGVFTNTIKLIVGR 72  
Db 90 TVPIWSVPVAVLLPI-IVFVCFYKRTCVYDLHHSILGLLFAVLITGVITDSIKVATGR 148  
Qy 73 PRDFFVRCPPDGL-----AHSIDLCTGDKDVVNEGRKSPFSGHSSFAFAGLAFASFYLAG 128  
Db 149 PRPNFYWRCPDGGKELYDALGGVVCHGKAAEVKEGHSFPSPGHTSWSFAGLTFLSLYLSG 208  
Qy 129 KLHCFTPQGRKSWRFCAFLSPLLFAAVIALSRCTCDYKHHQDL 172  
Db 209 KIKAFNNEGHVA--KLCVLIFPPLAACLGVISRVDYVHHWQDV 250  
Search completed: August 14, 2003, 13:07:23  
Job time : 96 secs



PT proteins, useful in gene therapy and vaccination against a variety of  
XX diseases, and for diagnosis of those diseases -  
PS Claim 21; Page 882-883; 921pp; English.  
XX The invention relates to full length GENSET human nucleic acids encoding  
CC potentially secreted proteins. The nucleic acids and the polypeptides  
CC they encode may be used in the prevention, treatment and diagnosis of  
CC diseases associated with inappropriate GENSET gene expression. For  
CC example, they be used to treat disorders associated with decreased  
CC GENSET gene expression by rectifying mutations or deletions in a  
CC patient's genome that affect the activity of GENSET or by supplementing  
CC the patient's own production of GENSET polypeptides. Conversely,  
CC antisense nucleic acid molecules may be administered to down regulate  
CC GENSET expression by binding with the cells' own genes and preventing  
CC their expression. The sense and antisense nucleic acids may also be  
CC used as DNA probes in diagnostic assays to detect and quantitate the  
CC presence of similar nucleic acid sequences in samples, and hence to  
CC determine which patients may be in need of restorative therapy.  
CC The GENSET polypeptides may be used as antigens in the production of  
CC antibodies and in assays to identify modulators (agonists and  
CC antagonists) of GENSET polypeptide expression and activity. The  
CC present sequence is a GENSET polypeptide of the invention.

XX Sequence 180 AA;  
SQ Query Match 100.0%; Score 968; DB 22; Length 180;  
Best Local Similarity 100.0%; Pred. No. 3.1e-110;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MWLYRNPVVEAEYFPTKPMFVIAFLSPLSLIFLAKFLKADTRDSRQACLAASLALALNG 60  
DB 1 MWLYRNPVVEAEYFPTKPMFVIAFLSPLSLIFLAKFLKADTRDSRQACLAASLALALNG 60  
QY 61 VFTNTIKLIVGRPRDPFFVRCPPDGLAHSDDLCTGDKDVVNEGRKSPSGHSSFAFAGLA 120  
DB 61 VFTNTIKLIVGRPRDPFFVRCPPDGLAHSDDLCTGDKDVVNEGRKSPSGHSSFAFAGLA 120  
QY 121 FASFYLAGKLHCFPTQGRGKSWRFCAFLSPLIFLAKFLKADTRDSRQACLAASLALALNG 180  
DB 121 FASFYLAGKLHCFPTQGRGKSWRFCAFLSPLIFLAKFLKADTRDSRQACLAASLALALNG 180

RESULT 2  
AAB70690  
ID AAB70690 standard; Protein; 175 AA.  
XX AAB70690;  
XX  
XX 17-MAY-2001 (first entry)  
XX Human hDPP protein sequence SEQ ID NO:7.  
XX Human; hDPP; diacylglycerol pyrophosphate phosphatase; DPP; detection.  
XX Homo sapiens.  
XX CN1271009-A.  
XX 25-OCT-2000.  
XX 17-MAR-2000; 2000CN-0114952.  
XX 17-MAR-2000; 2000CN-0114952.  
XX (SREN-) SOUTHERN RES CENT NAT HUMAN GENE GROUP.  
XX Li N, Xiao H, Liu F;  
XX WPI; 2001-081384/10.  
XX N-PSDB; AAF74766.  
XX New human diacyl glyceropyrophosphate phosphatase protein and its code

PT sequence -  
XX Claim 4; Page 17; 19pp; Chinese.  
XX The present invention describes a human diacylglycerol pyrophosphate  
CC phosphatase (DPP) designated hDPP. hDPP is expressed in normal tissue  
CC near cancerous liver cells of a human body. Also described are methods  
CC for the preparation and detection of hDPP nucleotide and protein  
CC sequences. The present sequence represents human hDPP, as given in the  
CC present invention.  
XX Sequence 175 AA;  
SQ Query Match 94.6%; Score 916; DB 22; Length 175;  
Best Local Similarity 98.3%; Pred. No. 7.1e-104;  
Matches 171; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 MWLYRNPVVEAEYFPTKPMFVIAFLSPLSLIFLAKFLKADTRDSRQACLAASLALALNG 60  
DB 1 MWLYRNPVVEAEYFPTKPMFVIAFLSPLSLIFLAKFLKADTRDSRQACLAASLALALNG 60  
QY 61 VFTNTIKLIVGRPRDPFFVRCPPDGLAHSDDLCTGDKDVVNEGRKSPSGHSSFAFAGLA 120  
DB 61 VFTNTIKLIVGRPRDPFFVRCPPDGLAHSDDLCTGDKDVVNEGRKSPSGHSSFAFAGLA 120  
QY 121 FASFYLAGKLHCFPTQGRGKSWRFCAFLSPLIFLAKFLKADTRDSRQACLAASLALALNG 174  
DB 121 FASFYLAGKLHCFPTQGRGKSWRFCAFLSPLIFLAKFLKADTRDSRQACLAASLALALNG 174  
RESULT 3  
AAM23571  
ID AAM23571 standard; Protein; 157 AA.  
XX AAM23571;  
XX 12-OCT-2001 (first entry)  
XX Arabidopsis EST encoded protein SEQ ID NO: 1096.  
XX Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;  
XX tomato; monkey; dog; sea urchin; expressed sequence tag; EST;  
XX diagnostics; forensic test; gene mapping; genetic disorder;  
XX biodiversity; gene therapy; nutrition.  
XX Arabidopsis thaliana.  
XX WO200154477-A2.  
XX 02-AUG-2001.  
XX 25-JAN-2001; 2001WO-US02687.  
XX 25-JAN-2000; 2000US-0491404.  
XX 17-JUL-2000; 2000US-0617746.  
XX 03-AUG-2000; 2000US-0631451.  
XX 15-SEP-2000; 2000US-0663870.  
XX (HYSE-) HYSEQ INC.  
XX Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;  
XX Cao Y, Drmanac RA, Zhang J, Werhman T;  
XX WPI; 2001-476164/51.  
XX N-PSDB; AAH98230.  
XX Isolated polypeptide for treatment of diseases, diagnostics, raising  
XX antibodies and research use -  
XX Claim 20; Page 821-822; 1275pp; English.  
XX The present invention provides the protein and coding sequences of novel  
XX proteins from a variety of organisms, including human, dog, cat, horse,  
XX



CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea  
CC urchin and tomato. These were derived from expressed sequence tags (ESTs)  
CC from the organism of interest. They can be used in diagnostics,  
CC forensics, gene mapping, identification of mutations, to assess  
CC biodiversity and for nutritional purposes. The present sequence is a  
CC protein of the invention.  
XX  
SQ Sequence 157 AA;

Query Match 83.6%; Score 809; DB 22; Length 157;  
Best Local Similarity 98.1%; Pred. No. 8.3e-91;  
Matches 153; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
Qy 19 MFVIAFLSPLSLIFLAKFLKADTRDSRQACLAASLALANGVFTNTIKLIVGRPRDPFF 78  
Db 1 MFVIAFLSPLSLIFLAKFLKADTRDSRQACLAASLALANGVFTNTIKLIVGRPRDPFF 60  
Qy 79 YRCFPDGLAHSDLMTGDKDVVNEGRKSFPSGHSFAGLAFASFYLAGKLHCFTPQGR 138  
Db 61 YRCFPDGLAHSDLMTGDKDVVNEGRKSFPSGHSFAGLAFASFYLAGKLHCFTPQGR 120  
Qy 139 GKSWRFCAFLSPLFAAVALSRCTCDYKHHWQDLK 174  
Db 121 GKSWRFCAFLSPLFAAVALSRCTCDYKHHWQGPFK 156

RESULT 4  
AAU23516  
ID AAU23516 standard; Protein; 233 AA.  
AC AAU23516;  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Novel human enzyme polypeptide #602.  
XX  
KW Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;  
KW ligase; hyperproliferative disorder; immunodeficiency disorder;  
KW autoimmune disorder; neurological disorder; metabolic disorder;  
KW inflammatory disorder; cardiovascular disorder; reproductive disorder;  
KW blood-related disorder; infectious disorder; cytostatic; anti arthritic;  
KW nephrotropic; anticoagulant.  
XX  
OS Homo sapiens.  
XX  
PN WO200155301-A2.  
XX  
XX  
PD 02-AUG-2001.

XX  
PF 17-JAN-2001; 2001WO-US01239.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226686.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 21-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.

CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 233 AA;  
Query Match 66.2%; Score 641; DB 22; Length 233;  
Best Local Similarity 66.7%; Pred. No. 5.7e-70;  
Matches 114; Conservative 27; Mismatches 30; Indels 0; Gaps 0;  
QY 1 MWLYRNPVVEAEYPTKPMFVIAFLSLJFLAKFLKADTRDSROACLAASLALALNG 60  
47 IWLKYNPLVQSDNIPTRLMFAISFLTPLAIVCVKLIIRRTDKTEIKFAFLAVSLALNG 106  
61 VFTNTIKLIVGRPRDPFFVRCPPDGLAHSDLCTGDKDVNVEGRKSPFSGHSSAFAGLA 120  
107 VCTNTIKLIVGRPRDPFFVRCPPDGVNSEMHCCTGDPDLVSEGRKSPFSSAFSGLG 166  
121 FASFLYLAGKLCFTTQGRKSWRFCAPLSPLFAAVIALSRCTCDYKHHWQD 171  
167 FTTFYLAGKLCFTTSGRKSRLCAAILFLYCAMMIALSRMCDYKHHWQD 217  
RESULT 5  
AAU21835  
ID AAU21835 standard; Protein; 233 AA.  
XX  
AC AAU21835;  
DT 06-DEC-2001 (first entry)  
DE Novel human neoplastic disease associated polypeptide #268.  
KW Human; neoplastic disease associated polypeptide; cancer;  
KW hyperproliferative disorder; neural disorder; immune system disorder;  
KW muscular disorder; reproductive disorder; gastrointestinal disorder;  
KW pulmonary disorder; cardiovascular disorder; renal disorder;  
KW neuroprotective; cytostatic; anti inflammatory; vasotropic.  
XX Homo sapiens.  
XX WO200155163-A1.  
XX  
XX 02-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US01358.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
XX 04-FEB-2000; 2000US-0180628.  
XX 24-FEB-2000; 2000US-0184664.  
XX 02-MAR-2000; 2000US-0186350.  
XX 16-MAR-2000; 2000US-0189874.  
XX 17-MAR-2000; 2000US-0190076.  
XX 18-APR-2000; 2000US-0198123.  
XX 19-MAY-2000; 2000US-0205515.  
XX 07-JUN-2000; 2000US-0209467.  
XX 28-JUN-2000; 2000US-0214886.  
XX 30-JUN-2000; 2000US-0215135.  
XX 07-JUL-2000; 2000US-0216647.  
XX 07-JUL-2000; 2000US-0216880.  
XX 11-JUL-2000; 2000US-0217487.  
XX 11-JUL-2000; 2000US-0217496.  
XX 14-JUL-2000; 2000US-0218290.  
XX 26-JUL-2000; 2000US-0220963.  
XX 26-JUL-2000; 2000US-0220964.  
XX 14-AUG-2000; 2000US-0224518.  
XX 14-AUG-2000; 2000US-0224519.  
XX 14-AUG-2000; 2000US-0225213.  
XX 14-AUG-2000; 2000US-0225214.  
XX 14-AUG-2000; 2000US-0225266.  
XX 14-AUG-2000; 2000US-0225267.  
XX 14-AUG-2000; 2000US-0225268.  
XX 14-AUG-2000; 2000US-0225270.  
XX 14-AUG-2000; 2000US-0225447.

08-NOV-2000; 2000US-0246526.  
08-NOV-2000; 2000US-0246526.  
08-NOV-2000; 2000US-0246527.  
08-NOV-2000; 2000US-0246528.  
08-NOV-2000; 2000US-0246532.  
08-NOV-2000; 2000US-0246609.  
08-NOV-2000; 2000US-0246610.  
08-NOV-2000; 2000US-0246611.  
08-NOV-2000; 2000US-0246613.  
08-NOV-2000; 2000US-0249207.  
17-NOV-2000; 2000US-0249208.  
17-NOV-2000; 2000US-0249209.  
17-NOV-2000; 2000US-0249210.  
17-NOV-2000; 2000US-0249211.  
17-NOV-2000; 2000US-0249212.  
17-NOV-2000; 2000US-0249213.  
17-NOV-2000; 2000US-0249214.  
17-NOV-2000; 2000US-0249215.  
17-NOV-2000; 2000US-0249216.  
17-NOV-2000; 2000US-0249217.  
17-NOV-2000; 2000US-0249218.  
17-NOV-2000; 2000US-0249244.  
17-NOV-2000; 2000US-0249245.  
17-NOV-2000; 2000US-0249264.  
17-NOV-2000; 2000US-0249265.  
17-NOV-2000; 2000US-0249297.  
17-NOV-2000; 2000US-0249299.  
17-NOV-2000; 2000US-0249300.  
01-DEC-2000; 2000US-0250160.  
01-DEC-2000; 2000US-0250391.  
05-DEC-2000; 2000US-0251030.  
05-DEC-2000; 2000US-0251988.  
05-DEC-2000; 2000US-0256719.  
06-DEC-2000; 2000US-0251479.  
08-DEC-2000; 2000US-0251856.  
08-DEC-2000; 2000US-0251868.  
08-DEC-2000; 2000US-0251869.  
08-DEC-2000; 2000US-0251989.  
08-DEC-2000; 2000US-0251990.  
11-DEC-2000; 2000US-0254097.  
05-JAN-2001; 2001US-0259678.  
(HUMA-) HUMAN GENOME SCI INC.  
Rosen CA, Barash SC, Ruben SM;  
WPI; 2001-465566/50.  
N-PSDB; AAS41386.  
Novel polypeptides and polynucleotides useful for diagnosing,  
preventing, treating neural, immune system, muscular, reproductive,  
pulmonary, cardiovascular, renal, proliferative disorders and cancerous  
diseases  
Claim 11; SEQ ID No 1512; 1180pp; English.  
The present invention relates to the isolation of novel human enzyme  
polypeptides, and the cDNA (AAS40785-AAS41684) and genomic sequences  
encoding them. The enzyme polypeptides of the invention may comprise the  
functional classes of oxidoreductases, transferases, hydrolases, lyases,  
isomerases or ligases. The sequences of the invention are useful in the  
diagnosis, treatment, prevention and/or prognosis of a wide range of  
disorders, including hyperproliferative disorders (e.g. cancer),  
immunodeficiency disorders (e.g. AIDS) autoimmune disorders  
(e.g. arthritis), neurological disorders (e.g. Alzheimer's disease),  
metabolic disorders (e.g. phenylketonuria), inflammatory disorders  
(e.g. asthma), cardiovascular disorders (e.g. atherosclerosis),  
blood-related disorders (e.g. haemophilia), reproductive disorders  
(e.g. infertility) and infectious disorders (e.g. Influenza). The  
polynucleotides of the invention can also be used in gene therapy.  
AAU22915-AAU23814 represent the novel human enzyme polypeptides of the  
invention.  
Note: The sequence data for this patent did not form part of the printed

PR 14-AUG-2000; 2000US-0225757.  
 PR 14-AUG-2000; 2000US-0225758.  
 PR 14-AUG-2000; 2000US-0225759.  
 PR 18-AUG-2000; 2000US-0226279.  
 PR 22-AUG-2000; 2000US-0226681.  
 PR 22-AUG-2000; 2000US-0226688.  
 PR 22-AUG-2000; 2000US-0227182.  
 PR 23-AUG-2000; 2000US-0227009.  
 PR 30-AUG-2000; 2000US-0228924.  
 PR 01-SEP-2000; 2000US-0229287.  
 PR 01-SEP-2000; 2000US-0229343.  
 PR 01-SEP-2000; 2000US-0229344.  
 PR 01-SEP-2000; 2000US-0229345.  
 PR 05-SEP-2000; 2000US-0229509.  
 PR 05-SEP-2000; 2000US-0229513.  
 PR 06-SEP-2000; 2000US-0230437.  
 PR 06-SEP-2000; 2000US-0230438.  
 PR 08-SEP-2000; 2000US-0231242.  
 PR 08-SEP-2000; 2000US-0231243.  
 PR 08-SEP-2000; 2000US-0231244.  
 PR 08-SEP-2000; 2000US-0231413.  
 PR 08-SEP-2000; 2000US-0231414.  
 PR 08-SEP-2000; 2000US-0232080.  
 PR 08-SEP-2000; 2000US-0232081.  
 PR 12-SEP-2000; 2000US-0231968.  
 PR 14-SEP-2000; 2000US-0232397.  
 PR 14-SEP-2000; 2000US-0232398.  
 PR 14-SEP-2000; 2000US-0232399.  
 PR 14-SEP-2000; 2000US-0232400.  
 PR 14-SEP-2000; 2000US-0232401.  
 PR 14-SEP-2000; 2000US-0233063.  
 PR 14-SEP-2000; 2000US-0233064.  
 PR 14-SEP-2000; 2000US-0233065.  
 PR 21-SEP-2000; 2000US-0234223.  
 PR 21-SEP-2000; 2000US-0234274.  
 PR 25-SEP-2000; 2000US-0234997.  
 PR 25-SEP-2000; 2000US-0234998.  
 PR 26-SEP-2000; 2000US-0235484.  
 PR 27-SEP-2000; 2000US-0235834.  
 PR 27-SEP-2000; 2000US-0235836.  
 PR 29-SEP-2000; 2000US-0236327.  
 PR 29-SEP-2000; 2000US-0236367.  
 PR 29-SEP-2000; 2000US-0236368.  
 PR 29-SEP-2000; 2000US-0236369.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0236802.  
 PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 02-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0239935.  
 PR 13-OCT-2000; 2000US-0239937.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 20-OCT-2000; 2000US-0241825.  
 PR 01-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.

PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249254.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 11-DEC-2000; 2000US-0251990.  
 PR 05-JAN-2001; 2001US-0255097.  
 PR 05-JAN-2001; 2001US-02559678.  
 XX  
 FA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX  
 DR WPI; 2001-465558/50.  
 DR N-PSDB; AAS35034.  
 XX  
 PT Novel polypeptides and polynucleotides useful as diagnostic reagents to  
 PT diagnose diseases or disorders associated with aberrant expression or  
 PT activity of polypeptides, and for treating cancers, rheumatoid  
 PT arthritis  
 XX  
 PS Claim 11; SEQ ID No 562; 687pp; English.

The present invention relates to the isolation of novel human neoplastic diseases associated polypeptides, and cDNA (AAS34767-AAS35050) and DNA sequences encoding for these polypeptides. The sequences of the invention are useful in the diagnosis, treatment, prevention and/or prognosis of disorders involving neoplastic disease such as hyperproliferative disorders (e.g. leukaemia, bone cancer, bladder cancer, brain stem glioma, adult liver cancer, childhood cerebellar astrocytoma, or Hodgkin's lymphoma). The sequences of the invention may also be useful for treating other disorders such as neural disorders, immune system disorders, muscular disorders, reproductive disorders, gastrointestinal disorders, pulmonary disorders, cardiovascular disorders and renal disorders. The polynucleotide sequences of the invention are also useful in gene therapy. AAU21568-AAU21851 represent the novel human neoplastic disease associated polypeptides of the invention.  
 Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 233 AA;

Query Match 66.2%; Score 641; DB 22; Length 233;  
 Best Local Similarity 66.7%; Pred. No. 5,7e-70;

[illegible]

CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell  
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),  
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,  
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal  
 CC vascular growth. Polypeptides involved with tissue regeneration and  
 CC repair (or nucleic acids encoding them) may be used to promote wound  
 CC healing (e.g., of burns, incisions and ulcers), while those with  
 CC immunomodulatory activities may be used in the treatment of viral,  
 CC bacterial and fungal infections in addition to immune disorders.  
 CC Polypeptides with growth factor activity may be used in cell cultures to  
 CC promote cell growth. For example, such polypeptides may be used to  
 CC manipulate stem cells in culture to give rise to neuroepithelial cells  
 CC that can be used to augment or replace cells damaged by illness,  
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides  
 CC may also be used in the diagnosis of the above conditions, and in drug  
 CC screening techniques. The present sequence represents a novel human  
 CC polypeptide of the invention.

XX Sequence 184 AA;

Query Match 56.6%; Score 548; DB 22; Length 184;  
 Best Local Similarity 64.9%; Pred. No. 1.1e-58;  
 Matches 98; Conservative 26; Mismatches 27; Indels 0; Gaps 0;

Qy 1 MWLYRNPVVEAEYPTKPMFVIAFLSLFLAKFLKADTRDSRQACLAALALNG 60  
 Db 34 IWLKYNPLGQSDNIPTRLMFAISLTLPLAVICVVKIIRRTDKTEKFAFLAVSLALALNG 93  
 Qy 61 VFTNTIKLVGRPRDPFYRCPPDGLAHSIDLCTGDKDVNVEGRKSPSGHSSFAFAGLA 120  
 Db 94 VCTNTIKLVGRPRDPFYRCPPDGVNSENHCITGDDPLVSEGRKSPSIHSSFAFSGLG 153  
 Qy 121 FASYLAGKLHCFTPOGRGKSWRFCAFLSPL 151  
 Db 154 FTTEYLAGKLHCFTESGRGKSWRLCAALPL 184

RESULT 8  
 ABB59271  
 ID ABB59271 standard; Protein; 363 AA.  
 XX  
 AC ABB59271;  
 XX  
 DT 26-MAR-2002 (first entry)  
 DE Drosophila melanogaster polypeptide SEQ ID NO 4605.  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 XX Drosophila melanogaster.  
 OS  
 XX WO200171042-A2.  
 PN  
 XX 27-SEP-2001.  
 PD  
 XX 23-MAR-2001; 2001WO-US09231.  
 PF  
 XX 23-MAR-2000; 2000US-191637P.  
 PR  
 XX 11-JUL-2000; 2000US-0614150.  
 XX (PEKE ) PE CORP NY.  
 PA  
 XX Venter JC, Adams M, Li PWD, Myers EW;  
 PI  
 XX WPI; 2001-656860/75.  
 DR  
 XX N-PSDB; ABL03374.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX  
 PS Disclosure; SEQ ID NO 4605; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins  
 CC (ABB5737-ABB72072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 363 AA;

Query Match 45.5%; Score 440; DB 22; Length 363;  
 Best Local Similarity 44.5%; Pred. No. 4.7e-45;  
 Matches 85; Conservative 29; Mismatches 49; Indels 28; Gaps 4;

Qy 1 MWLYRNPVVEAEYPTKPMFVIAFLSLFLAKFLKADTRDSRQACLAAS 53  
 Db 103 LWLYKN-----PRRPDIVRGGELLFWIVVAPF-LVTIAFYWYTRDRDRFRAASWAWT 153  
 Qy 54 LALALNGVFTNTIKLVGRPRDPFYRCPPDGL-----AHSIDLCTGDKOVVN 101  
 Db 154 LALCMNGIPTSVLKITVGRPRDPFYRCPPDGVNVLNTTNSGVDTSLDFNCTGLPGDIN 213  
 Qy 102 EGRKSPSGHSSFAFAGLAFAFSLAGKLHCFTPOGRGKSWRFCAFLSPLFAAVIALSR 161  
 Db 214 EGRKSPSGHSSFAFAGLAFAFSLAGKLHCFTPOGRGKSWRFCAFLSPLFAAVIALSR 273  
 Qy 162 TCDYKHHWQDL 172  
 Db 274 TCDYKHHWQDV 284

RESULT 9  
 ABB59272  
 ID ABB59272 standard; Protein; 412 AA.  
 XX  
 AC ABB59272;  
 XX  
 DT 26-MAR-2002 (first entry)  
 DE Drosophila melanogaster polypeptide SEQ ID NO 4608.  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 XX Drosophila melanogaster.  
 OS  
 XX WO200171042-A2.  
 PN  
 XX 27-SEP-2001.  
 PD  
 XX 23-MAR-2001; 2001WO-US09231.  
 PF  
 XX 23-MAR-2000; 2000US-191637P.  
 PR  
 XX 11-JUL-2000; 2000US-0614150.  
 XX (PEKE ) PE CORP NY.  
 PA  
 XX Venter JC, Adams M, Li PWD, Myers EW;  
 PI  
 XX WPI; 2001-656860/75.  
 DR  
 XX N-PSDB; ABL03375.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX  
 PS Disclosure; SEQ ID NO 4608; 21pp + Sequence Listing; English.

ST  
CLAIM #; Page 100-101; Page 100-101

mammalian skin cells. The polypeptide is useful for stimulating keratinocyte growth and motility, inhibiting the growth of cancer cells, modulating angiogenesis, inhibiting angiogenesis and vascularisation of tumours, modulating skin inflammation, stimulating the growth of epithelial cells, inhibiting the binding of human immunodeficiency virus (HIV)-1 to leukocytes, and treating inflammatory disease, cancer and neurological diseases. The polynucleotide can be used as a marker, in the identification of genetic disorders, and for the design of oligonucleotides for examining expression patterns.

```

SQ Sequence 128 AA;
Query Match 40.2%; Score 389; DB 22; Length 128;
Best Local Similarity 89.0%; Pred. No. 2.1e-39;
Matches 73; Conservative 4; Mismatches 5; Indels 0; Gaps 0

Qy 1 MWLYRNPVVEAEYPTTQPMFVIAFLPSLILFLAKFLKKADTRDSRQACLAASLALALNG 60
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 47 LWLYRNPVVEAEYPTTQPMFVIAFLPSLILFLAKFLKKADTRDSRQACLAASLALALNG 106

Qy 61 VFTNTIKLIVGRPRDPFFYRCF 82
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 107 VFTNTIKLIVGRPRDPFFYRCF 128
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

RESULT 12	
ABB72200	
ID	ABB72200 standard; Protein; 128 AA.
XX	
XX	
XX	ABB72200;
XX	
XX	
DT	04-APR-2002 (first entry)
XX	
DE	Rat protein isolated from skin cells SEQ ID NO: 316.
DE	
XX	
XX	Human; rat; mouse; skin cell; skin wound; cancer; growth defect;
XX	developmental defect; inflammatory disease; dermatological; vulnery;
KW	immunomodulator; anti-inflammatory; cytostatic; neuroprotective.
KW	

AA	Rattus sp.
OS	
XX	WO200190357-A1.
XX	
XX	29-NOV-2001.
XX	
XX	24-MAY-2001; 2001WO-NZ00099.
XX	
XX	24-MAY-2000; 2000US-206650P.
PR	25-JUL-2000; 2000US-221232P.
XX	
XX	(GENE-) GENESIS RES & DEV CORP LTD.
PA	
XX	
XX	Watson JD, Strachan L, Sleeman M, Onrust R, Murison JG, Kumble KD;
PI	
XX	WPI; 2002-122020/16.
XX	
XX	New polynucleotides and polypeptides encoded by the polynucleotides
PT	isolated from skin cells, useful for treating skin wounds, cancers,
PT	growth and developmental defects, inflammatory diseases, or for
PT	modulating immune responses -
PT	

Example 2; Page 208-209; 466pp; English.

XX

```

SQ Sequence 128 AA;
Query Match 40.2%; Score 389; DB 23; Length 128;
Best Local Similarity 89.0%; Pred. No. 2.1e-39;
Matches 73; Conservative 4; Mismatches 5; Indels 0; Gaps 0

Qy 1 MWLYRNPYVEAEYFPTKPMFVIAFLSPLSLIIFLAKFLKKADTRDSRQACLAASLALALNG 60
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 47 LWLYRNPYVEAEYFPTGPMFVIAFLTSLTSLIIFFAKFLRKADATDSKQACLAASLALALNG 106
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 61 VFTNTIKLIVGRPRDPFFYRCF 82
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 107 VFTNIIKLIVGRPRDPFFYRCF 128
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 13
AAG05044
ID AAG05044 standard; Protein; 292 AA.
XX AC
XX AAG05044;
XX AC
XX XX
XX 17-OCT-2000 (first entry)
XX DT
XX DE
XX DE
XX XX
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; Genetic mapping; gene expression control; promoter;
XX KW
XX KW termination sequence.

```

OS *Arabidopsis thaliana*.

PN EP1033405-A2.

06-SEP-2000  
PD  
YYXX  
DE FEB 2000

XX

PR 05-MAR-1999; 99US-0123180.

PR 23-MAR-1999; 99US-0125788.

23-MAR-1999, 3303-0120204  
29-MAR-1999: 99US-0126785

PR 01-APR-1999;  
99US-0127462;  
PR 06-APR-1999;  
99US-0128234

PR 08-APR-1999; 99US-0128714.

PR 19-APR-1999; 99US-0130077.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0131449.

30-APR-1999: 99US-0132407.

PR 04-MAY-1999; 99US-0132484;  
PP 05-MAY-1999; 99US-0132485

PR 06-MAY-1999; 99US-0132486.  
 DE 06 MAY 1999 00132487  
 DE 06 MAY 1999 00132487

PR 07-MAY-1999; 99US-0132863.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134221.

18-MAY-1999: 99US-0134768

PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999. 99US-0135124

PR 21-MAY-1999; 99US-0135353.  
 DD 24 MAY 1999; 00RIC 0135353.  
 DD 24 MAY 1999; 00RIC 0135353.

PR 25-MAY-1999; 99US-0136021.

PR 28-MAY-1999; 99US-0136782.

2  
2  
2  
-  
0  
4  
5  
  
0  
0  
1  
1  
  
-  
1  
1  
1  
1  
4  
0  
0  
  
1  
0  
4

Thu Apr 21 12:12:18 2005

us-09-876-997-399.rag

PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144086.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144333.  
PR 20-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0146389.  
PR 04-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147392.

PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149930.  
PR 23-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 13-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159594.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match

39.5%; Score 382; DB 21; Length 292;





```

PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 08-AUG-1999; 99US-0147303.
PR 08-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 39.5%; Score 382; DB 21; Length 292;
Best Local Similarity 47.0%; Pred. NO. 4.6e-38;
Matches 77; Conservative 28; Mismatches 49; Indels 10; Gaps 4;
QY 16 TKPMF---VIAFLSPSLIFLAKFKADTRDSROACLAASLALALNGVFTNTIKLIVGR 72
Db 43 TVPIWSVPVYAMLLPL-VIFIFIYFRRDYDLDHHAIVLGLLYSLVLTAVLTDKNAVGR 101
QY 73 PREDFVRCFPDGLA---HSDLMCTGDKDVNNEGRKSPSGHSSFAFAGLAFASFYLAG 128
Db 102 PREDFVRCFPDGLA---HSDLMCTGDKDVNNEGRKSPSGHSSFAFAGLAFASFYLAG 128
QY 129 KLHCFTPQGRGKSWRFCAFLSPLLFAAVIALSRCTCYKHHWQDL 172
Db 162 KIQAF--DGKGVAKLCIVILPLLFAALVGISRVDYVHHWQDV 203

RESULT 15
AAG14938
ID AAG14938 standard; Protein; 292 AA.
XX AAG14938;
XX AAG14938;
DT 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 14988.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
XX EP1033405-A2.
PD 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 11-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.

```

PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

```
Query Match      39.5%; Score 382; DB 21; Length 292;
Best Local Similarity 47.0%; Pred. No. 4.6e-38;
Matches 77; Conservative 28; Mismatches 49; Indels 10; Gaps 4;

Qy 16 TKPMF---VIAFLSPLSLIFLAKFLKADTRDSRQACLAASLALALNGVFTNTIKLIVGR 72
   |::| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 43 TVPIWSVVPVYAMLPL-VIFIFIYFRRRDVYDLHHAVLGLIYSVLVTAVLTDALKNAVGR 101
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 73 PRPDFFYRCPPDGLA---HSDLMCTGDKOVNKGKSPSPGHSSFAFAGLAFASFYLAG 128
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 102 PRPDFFWRCPPDGLKALYDSLGDVICHGDKSVIREGHKSPSPGHSTWSFSGLGFLSLYLSG 161
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 129 KLHCFTPGGRGKSWRFCAFLSPLIFAAVIALSRTCDYKHHWQDL 172
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 162 KIQAF--DGKGVAKLCIVILPLLFAALVGISRVDYWHHWQDV 203
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: August 14, 2003, 13:05:41  
Job time : 87 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 14, 2003, 13:03:37 ; Search time 30 Seconds  
(without alignments)  
253.865 Million cell updates/sec

Title: US-09-876-997-399  
Perfect score: 968  
Sequence: 1 MWLYRNPYVEAEYFPPTKPMF.....RTCDYKHQDILLKCTNTAK 180

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PTUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	389	40.2	128	3	US-09-188-930-316
2	389	40.2	128	4	US-09-312-283C-316
3	382	39.5	314	4	US-09-360-376-14
4	377	38.9	322	4	US-09-360-376-16
5	372	38.4	290	4	US-09-360-376-12
6	357	36.9	343	4	US-09-360-376-17
7	337	34.8	348	4	US-09-360-376-13
8	335	34.6	310	4	US-09-360-376-15
9	247	25.5	289	4	US-09-360-376-52
10	203.5	21.0	123	3	US-09-188-930-170
11	203.5	21.0	123	4	US-09-312-283C-170
12	184.5	19.1	282	4	US-09-360-376-54
13	183.5	19.0	283	3	US-08-992-035A-3
14	183.5	19.0	283	4	US-09-360-376-53
15	174.5	18.0	285	3	US-08-992-035A-1
16	127	13.1	233	4	US-09-360-376-55
17	114	11.8	323	4	US-09-122-315C-18
18	80	8.3	199	4	US-09-107-532A-7261
19	76	7.9	244	4	US-09-252-991A-21937
20	74	7.6	214	4	US-09-328-352-4246
21	72.5	7.5	401	4	US-09-252-991A-27198
22	72.5	7.5	1765	4	US-09-354-147C-5
23	72	7.4	444	4	US-09-328-352-5249
24	71.5	7.4	1765	4	US-09-354-147C-2
25	71.5	7.4	1765	4	US-09-354-147C-3
26	70.5	7.3	254	4	US-09-107-532A-4025
27	70	7.2	202	4	US-09-252-991A-26505

28	70	7.2	235	3	US-09-248-335-64	Sequence 64, Appl
29	69	7.1	1684	3	US-08-665-259-25	Sequence 25, Appl
30	69	7.1	1684	3	US-08-762-500-25	Sequence 25, Appl
31	69	7.1	1704	3	US-08-762-500-75	Sequence 75, Appl
32	69	7.1	2233	2	US-08-569-853-1	Sequence 1, Appli
33	69	7.1	2233	2	US-08-569-853-2	Sequence 2, Appli
34	69	7.1	2233	3	US-08-987-439-1	Sequence 1, Appli
35	67.5	7.0	429	4	US-09-351-150A-5	Sequence 5, Appli
36	67.5	7.0	432	4	US-09-252-991A-22739	Sequence 22739, A
37	67.5	7.0	444	4	US-08-482-746-15	Sequence 15, Appl
38	67.5	7.0	444	4	US-09-580-734-15	Sequence 15, Appl
39	67.5	7.0	444	4	US-08-374-009-15	Sequence 15, Appl
40	67	6.9	308	4	US-09-328-352-4190	Sequence 4190, Ap
41	67	6.9	940	4	US-09-198-452A-500	Sequence 500, App
42	66.5	6.9	233	4	US-09-134-001C-4013	Sequence 4013, Ap
43	65.5	6.8	315	4	US-09-674-529B-15	Sequence 15, Appl
44	65.5	6.8	1968	1	US-07-745-206A-7	Sequence 7, Appli
45	65.5	6.8	1968	1	US-08-455-543A-45	Sequence 45, Appl

ALIGNMENTS

RESULT 1  
US-09-188-930-316  
; Sequence 316, Application US/09188930A  
; Patent No. 6150502  
; GENERAL INFORMATION:  
; APPLICANT: Watson, James D.  
; APPLICANT: Strachan, Lorna  
; APPLICANT: Sleeman, Matthew  
; APPLICANT: Onrust, Rene  
; APPLICANT: Murison, James Greg  
; TITLE OF INVENTION: Compositions Isolated From Skin Cells  
; TITLE OF INVENTION: and Methods For Their Use  
; FILE REFERENCE: 11000.1011c1  
; CURRENT APPLICATION NUMBER: US/09/188,930A  
; CURRENT FILING DATE: 1998-11-09  
; NUMBER OF SEQ ID NOS: 348  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 316  
; LENGTH: 128  
; TYPE: PRT  
; ORGANISM: Rat  
US-09-188-930-316

Query Match	40.2%	Score 389	DB 3	Length 128
Best Local Similarity	89.0%	Pred. No. 7e-43	5	Indels 0
Matches	73	Conservative	4	Mismatches 0
Gaps	0			
QY	1	MWLYRNPYVEAEYFPPTKPMFVIAFLSPSLIFLAKFLKADTRDSROACLAASLALANG	60	
Db	47	LWLYRNPYVEAEYFPPTKPMFVIAFLSPSLIFLAKFLKADTRDSROACLAASLALANG	106	
QY	61	VFTNTKLVGRPRDPFFYRCF	82	
Db	107	VFTNTKLVGRPRDPFFYRCF	128	

RESULT 2  
US-09-312-283C-316  
; Sequence 316, Application US/09312283C  
; Patent No. 6573095  
; GENERAL INFORMATION:  
; APPLICANT: Watson, James D.  
; APPLICANT: Strachan, Lorna  
; APPLICANT: Sleeman, Matthew  
; APPLICANT: Onrust, Rene  
; APPLICANT: Murison, James G.  
; APPLICANT: Kumble, Krishanand D.  
; TITLE OF INVENTION: Compositions Isolated from Skin Cells  
; TITLE OF INVENTION: and Methods for Their Use  
; FILE REFERENCE: 11000.1011c2

```

; CURRENT APPLICATION NUMBER: US/09/312,283C
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 425
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 316
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Mouse
US-09-312-283C-316

Query Match 40.2%; Score 389; DB 4; Length 128;
Best Local Similarity 89.0%; Pred. No. 7e-43;
Matches 73; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 MWLYRNPVEAEYPTKPMFVIAFLPSLSLIIFLAKFLKADTRDSQACLAASLALALNG 60
Db 47 LWLYRNPVEAEYPTKPMFVIAFLTPLSLIFFAKFLRKADATDSKQACLAASLALNG 106
QY 61 VFTNIKLIVGRPRDPFFYRCF 82
Db 107 VFTNIKLIVGRPRDPFFYRCF 128

RESULT 3
US-09-360-376-14
; Sequence 14, Application US/09360376
; Patent No. 6495739
; GENERAL INFORMATION:
; APPLICANT: Lassar, Michael
; APPLICANT: Ruezinsky, Diane
; TITLE OF INVENTION: PLANT PHOSPHATIDIC ACID PHOSPHATASES
; FILE REFERENCE: 17026/01/US
; CURRENT APPLICATION NUMBER: US/09/360,376
; CURRENT FILING DATE: 1999-07-23
; PRIOR APPLICATION NUMBER: US 09/122,315
; PRIOR FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Arabidopsis sp.
US-09-360-376-14

Query Match 39.5%; Score 382; DB 4; Length 314;
Best Local Similarity 47.0%; Pred. No. 2e-41;
Matches 77; Conservative 28; Mismatches 49; Indels 10; Gaps 4

QY 16 TKPMF---VIAFLPSLSLIIFLAKFLKADTRDSQACLAASLALALNGVFTNTIKLIVGR 72
Db 65 TVPIWSVPVYAMLLPL-VIFIPIYFRRRDYYDLHVAVLGLLYSVLTAVLTDAIKNAVGR 123
QY 73 PRDPRFYRCFPDGLA---HSDLMCTGDKDVVNEGRKSPFGSHSSFAFAGLAFASFYLAG 128
Db 124 PRDPRFYRCFPDGLA---HSDLMCTGDKDVVNEGRKSPFGSHSSFAFAGLAFASFYLAG 183
QY 129 KLHCFTPOGRGKSWRFCAFLSPLFAAVIALSRDCTYKHHQDL 172
Db 184 KIQAF--DGKGHVAKLCIVILPLLFAALVIGISRVDDYWHHQQDV 225

RESULT 4
US-09-360-376-16
; Sequence 16, Application US/09360376
; Patent No. 6495739
; GENERAL INFORMATION:
; APPLICANT: Lassar, Michael
; APPLICANT: Ruezinsky, Diane
; TITLE OF INVENTION: PLANT PHOSPHATIDIC ACID PHOSPHATASES
; FILE REFERENCE: 17026/01/US
; CURRENT APPLICATION NUMBER: US/09/360,376
; CURRENT FILING DATE: 1999-07-23
; CURRENT APPLICATION NUMBER: US 09/122,315

```



1 MWLYRNPYVEAEYFFPTKPMFVIAFLSPSLIFLAKFLKKADTRDSK-----QACDASGE 31



```
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/992.035A
; FILING DATE: December 17, 1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0433 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 283 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: 1487873
; US-08-992-035A-3

Query Match 19.0%; Score 183.5; DB 3; Length 283;
Best Local Similarity 31.9%; Pred. No. 1.6e-15;
Matches 59; Conservative 17; Mismatches 54; Indels 55; Gaps 8;

QY 3 LYRNPVYEAFFPKPMFVIAFLSPLSLIFLAKFLKKADTRDSRQACLAASLALANGVF 62
DB 84 LHNSFVGNPYIATYKAVGAFLGVSS-----ASQSL----- 115

QY 63 TNTIKLIVGRPRDPFFYRCFP-----DGLAHSIDLMTCTGDKDVVNEGRKSPSPGSHSF 114
DB 116 TDIKTYTIGSLRPHFLAICNPDWKSKNCSDGYI-EDYICQNEBKVEGRSLSFYSGHSSF 174

QY 115 AFAGLAFASFYLAGKLHCFTFQGRGK-SWRFCAPLSPL-----FAAVTALSRTCDYKH 167
DB 175 SMYCMFLVALYL-----QARMKGDW--ARLLRPMLOFGLIAFSIYVGLSRVSDYKH 223

QY 168 HWODL 172
DB 224 HWSVDV 228

RESULT 14
US-09-360-376-53
; Sequence 53, Application US/09360376
; Patent No. 8495739
; GENERAL INFORMATION:
; APPLICANT: Lassar, Michael
; APPLICANT: Ruzizinsky, Diane
; TITLE OF INVENTION: PLANT PHOSPHATIDIC ACID PHOSPHATASES
; FILE REFERENCE: 17026/01/US
; CURRENT APPLICATION NUMBER: US/09/360.376
; CURRENT FILING DATE: 1999-07-23
; PRIOR APPLICATION NUMBER: US 09/122.315
; PRIOR FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 53
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-360-376-53

Query Match 19.0%; Score 183.5; DB 4; Length 283;
Best Local Similarity 31.9%; Pred. No. 1.6e-15;
Matches 59; Conservative 17; Mismatches 54; Indels 55; Gaps 8;

QY 3 LYRNPVYEAFFPKPMFVIAFLSPLSLIFLAKFLKKADTRDSRQACLAASLALANGVF 62
DB 84 LHNSFVGNPYIATYKAVGAFLGVSS-----ASQSL----- 115
```

```
QY 63 TNTIKLIVGRPRDPFFYRCFP-----DGLAHSIDLMTCTGDKDVVNEGRKSPSPGSHSF 114
DB 116 TDIKTYTIGSLRPHFLAICNPDWKSKNCSDGYI-EDYICQNEBKVEGRSLSFYSGHSSF 174

QY 115 AFAGLAFASFYLAGKLHCFTFQGRGK-SWRFCAPLSPL-----FAAVTALSRTCDYKH 167
DB 175 SMYCMFLVALYL-----QARMKGDW--ARLLRPMLOFGLIAFSIYVGLSRVSDYKH 223

QY 168 HWODL 172
DB 224 HWSVDV 228

RESULT 15
US-08-992-035A-1
; Sequence 1, Application US/08992035A
; Patent No. 6242179
; GENERAL INFORMATION:
; APPLICANT: Shah, Purvi
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: HUMAN PHOSPHATASES
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq For Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/992.035A
; FILING DATE: December 17, 1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0433 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 285 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BLADNOT06
; CLONE: 1719418
; US-08-992-035A-1

Query Match 18.0%; Score 174.5; DB 3; Length 285;
Best Local Similarity 30.9%; Pred. No. 2.4e-14;
Matches 60; Conservative 25; Mismatches 70; Indels 39; Gaps 9;

QY 7 PYEAEYFPKPMFVIAFLSPLSLIFLAKFL-----KKADTRDSRQACLAASLAL 58
DB 49 PYHDSVTST-VLILVGVGLPISIIIGETLSVYCNLLHNSFRNNYIATYKAITGFL 107

QY 59 NGV-----FNTTKLIVGRPRDPFFYRCFPD--GLAHS-----LMCTGDKDVVNEGRKS 106
DB 108 FGAASQSLTDIAKYSIGRLRPHFLDVCDDPDKSKINCSDGYIEYICRGAERKVEGRSL 167
```

us-09-876-997-399.rai

Thu Apr 21 12:12:18 2005

Qy 107 FPSGHSFAFAGLAFAFYLAGKLHCFTPGRGK-SWRFCAPLSPLLFAAVIA-----L 159  
Db 168 FYSGHSFSFMYCMLFVALYL-----QARMKGDW--ARLLRPTLOFGLVAVSIYVGL 216  
Qy 160 SRTCDYKHHWQDLL 173  
Db 217 SRVSDYKHHWSDVL 230

Search completed: August 14, 2003, 13:07:59  
Job time : 31 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 14, 2003, 13:07:30 ; Search time 56 Seconds  
(without alignments)  
421.082 Million cell updates/sec

Title: US-09-876-997-399

Perfect score: 968

Sequence: 1 MWLYRNPYVEAEYFPKPMF.....RTCDYKHHWODLLKCTNTAK 180

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 492763 seqs, 131003257 residues

Total number of hits satisfying chosen parameters: 492763

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	968	100.0	180	10	US-09-731-872-399
2	641	66.2	233	15	Sequence 399, App
3	633	65.4	235	12	Sequence 562, App
4	389	40.2	128	11	Sequence 99, Appl
5	329	34.0	149	15	Sequence 316, App
6	203.5	21.0	123	11	Sequence 452, App
7	138	14.3	220	15	Sequence 170, App
8	115	11.9	180	10	Sequence 5750, App
9	115	11.9	180	15	Sequence 125, App
10	115	11.9	183	15	Sequence 490, App
11	110	11.4	203	15	Sequence 340, App
12	93	9.6	525	15	Sequence 493, App
13	87	9.0	84	9	Sequence 15017, A
14	82	8.5	497	15	Sequence 40341, A
15	77.5	8.0	184	9	Sequence 9418, App
					Sequence 860, App

16	77.5	8.0	184	11	US-09-925-299-860	Sequence 860, App
17	73.5	7.6	283	15	US-10-156-761-8708	Sequence 8708, Ap
18	73	7.5	216	9	US-09-815-242-13375	Sequence 13375, A
19	72	7.4	580	15	US-10-004-551-10	Sequence 10, Appl
20	70.5	7.3	159	10	US-09-738-626-6679	Sequence 6679, Ap
21	68	7.0	276	9	US-09-741-669-302	Sequence 302, App
22	67.5	7.0	444	9	US-09-853-386-132	Sequence 132, App
23	67.5	7.0	444	9	US-09-853-386-136	Sequence 136, App
24	67.5	7.0	444	9	US-09-191-724-15	Sequence 15, Appl
25	67.5	7.0	444	15	US-10-242-822B-2	Sequence 2, Appli
26	67	6.9	410	15	US-10-156-761-13608	Sequence 13608, A
27	66.5	6.9	227	8	US-08-834-705-16	Sequence 15, Appl
28	66.5	6.9	227	9	US-09-815-242-11375	Sequence 11375, A
29	66.5	6.9	228	9	US-09-815-242-11342	Sequence 11542, A
30	66.5	6.9	274	15	US-10-156-761-10121	Sequence 10121, A
31	66.5	6.9	705	11	US-09-907-907A-43	Sequence 43, Appl
32	65.5	6.8	296	10	US-09-771-730-69	Sequence 69, Appl
33	65.5	6.8	296	10	US-09-771-730-116	Sequence 116, App
34	65.5	6.8	296	10	US-09-771-730-120	Sequence 120, App
35	65.5	6.8	303	10	US-09-771-730-118	Sequence 118, App
36	65.5	6.8	332	10	US-09-886-055-63	Sequence 63, Appl
37	65.5	6.8	332	10	US-09-771-730-10	Sequence 10, Appl
38	65.5	6.8	332	10	US-09-771-730-28	Sequence 28, Appl
39	65.5	6.8	332	11	US-09-804-291-63	Sequence 63, Appl
40	65.5	6.8	332	12	US-10-017-161-244	Sequence 244, App
41	65.5	6.8	332	15	US-10-236-250-2	Sequence 2, Appli
42	65.5	6.8	380	9	US-09-815-242-11291	Sequence 11291, A
43	65.5	6.8	383	12	US-10-259-165-202	Sequence 202, App
44	65.5	6.8	453	15	US-10-220-382-19	Sequence 19, Appl
45	65	6.7	320	15	US-10-242-822B-32	Sequence 32, Appl

ALIGNMENTS

RESULT 1  
US-09-731-872-399  
; Sequence 399, Application US/09731872  
; Patent No. US20020102604A1  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, Jean Baptiste  
; APPLICANT: Bougueleret, Lydie  
; APPLICANT: Jobert, Severin  
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS  
; FILE REFERENCE: 78.US3.REG  
; CURRENT APPLICATION NUMBER: US/09/731,872  
; CURRENT FILING DATE: 2000-12-07  
; PRIOR APPLICATION NUMBER: US 60/169,629  
; PRIOR FILING DATE: 1999-12-08  
; PRIOR APPLICATION NUMBER: US 60/187,470  
; PRIOR FILING DATE: 2000-03-06  
; NUMBER OF SEQ ID NOS: 482  
; SOFTWARE: Patent.pm  
; SEQ ID NO 399  
; LENGTH: 180  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-731-872-399

Query Match	100.0%;	Score 968;	DB 10;	Length 180;
Best Local Similarity	100.0%;	Pred. No. 1.3e-108;		
Matches 180;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MWLYRNPYVEAEYFPKPMFVIAFLPSLSLIFLAKFLKADTRDSROACLAASLALALNG	60	
Db	1	MWLYRNPYVEAEYFPKPMFVIAFLPSLSLIFLAKFLKADTRDSROACLAASLALALNG	60	
Qy	61	VFTNTIKLIVGRPRPFFYRCFDPGLAHSJDMCTGDKVNVNEGRKSPFSGHSSFAFAGLA	120	
Db	61	VFTNTIKLIVGRPRPFFYRCFDPGLAHSJDMCTGDKVNVNEGRKSPFSGHSSFAFAGLA	120	
Qy	121	FASFYLAGKLUHCTPOGRGKSWRFCAFLPSLLFAAVIALSRCTDYKHHWODLLKCTNTAK	180	

Db 121 FASFLAGKLCFTPOGRGKSWRFCAFLSPLLFAAVIALSRCTDYKHHWODLLKCTNTAK 180

RESULT 2

US-10-103-313-562

Sequence 562, Application US/10103313

Publication No. US20030082758A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: P0207C1

CURRENT APPLICATION NUMBER: US/10/103,313

CURRENT FILING DATE: 2002-03-12

NUMBER OF SEQ ID NOS: 653

Prior Application removed - See File Wrapper or Palm

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 562

LENGTH: 233

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (14)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-10-103-313-562

Query Match 66.2%; Score 641; DB 15; Length 233;

Best Local Similarity 66.7%; Pred. No. 5.1e-69;

Matches 114; Conservative 27; Mismatches 30; Indels 0; Gaps 0;

Qy 1 MWLYRNPYVEAEYPTKPMFVIAFLSPLSLFLAKFLKADTRDSRQACLAASLALALNG 60

Db 47 IWLKYNPLVQSDNITRLMFAISFLTPLAIVCVVKIIRRTDKTEIKFAFLAVSLALALNG 106

Qy 61 VFTNTIKLIVGRPRDPFFYRCFPGDLAHSMDLCTGDKDVNVEGRKSPFSGHSSFAFAGLA 120

Db 107 VCTNTIKLIVGRPRDPFFYRCFPGDVNVEGRKSPFSGHSSFAFAGLA 166

Qy 121 FASFLAGKLCFTPOGRGKSWRFCAFLSPLLFAAVIALSRCTDYKHHWOD 171

Db 167 FTFYLAGKLCFTSGRGKSWRLCAAILPLYCAMMIALSRMCDYKHHWOD 217

RESULT 3

US-10-241-220-99

Sequence 99, Application US/10241220

Publication No. US20030148408A1

GENERAL INFORMATION:

APPLICANT: Franz, Gretchen

APPLICANT: Hillan, Kenneth J.

APPLICANT: Phillips, Heidi

APPLICANT: Polakis, Paul

APPLICANT: Spencer, Susan

APPLICANT: Williams, P. Mickey

APPLICANT: Wu, Thomas

APPLICANT: Zhang, Zemin

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND TREATMENT OF TUMOR

FILE REFERENCE: P5010R1-US

CURRENT APPLICATION NUMBER: US/10/241,220

CURRENT FILING DATE: 2002-12-13

NUMBER OF SEQ ID NOS: 120

SEQ ID NO 99

LENGTH: 235

TYPE: PRT

ORGANISM: Homo Sapien

US-10-241-220-99

Query Match 65.4%; Score 633; DB 12; Length 235;

Best Local Similarity 66.4%; Pred. No. 4.8e-68;

Matches 113; Conservative 27; Mismatches 31; Indels 0; Gaps 0;

Qy 1 MWLYRNPYVEAEYPTKPMFVIAFLSPLSLFLAKFLKADTRDSRQACLAASLALALNG 60

Db 36 IWLKYNPLVQSDNITRLMFAISFLTPLAIVCVVKIIRRTDKTEIKFAFLAVSLALALNG 95

Qy 61 VFTNTIKLIVGRPRDPFFYRCFPGDLAHSMDLCTGDKDVNVEGRKSPFSGHSSFAFAGLA 120

Db 96 VCTNTIKLIVGRPRADFFYRCFPGDVNVEGRKSPFSGHSSFAFAGLA 155

Qy 121 FASFLAGKLCFTPOGRGKSWRFCAFLSPLLFAAVIALSRCTDYKHHWOD 171

Db 156 FTFYLAGKLCFTSGRGKSWRLCAAILPLYCAMMIALSRMCDYKHHWOD 206

RESULT 4

US-09-866-050A-316

Sequence 316, Application US/09866050A

Publication No. US20030040471A1

GENERAL INFORMATION:

APPLICANT: Watson, James D.

APPLICANT: Strachan, Iorna

APPLICANT: Sleeman, Matthew

APPLICANT: Orrust, Rene

APPLICANT: Murison, James G.

APPLICANT: Kumble, Krishanand D.

TITLE OF INVENTION: Compositions Isolated From Skin Cells

TITLE OF INVENTION: and Methods for Their Use

FILE REFERENCE: 11000.1011c4U

CURRENT APPLICATION NUMBER: US/09/866,050A

CURRENT FILING DATE: 2001-05-24

NUMBER OF SEQ ID NOS: 725

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 316

LENGTH: 128

TYPE: PRT

ORGANISM: Rat

US-09-866-050A-316

Query Match 40.2%; Score 389; DB 11; Length 128;

Best Local Similarity 89.0%; Pred. No. 5.8e-39;

Matches 73; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 MWLYRNPYVEAEYPTKPMFVIAFLSPLSLFLAKFLKADTRDSRQACLAASLALALNG 60

Db 47 IWLKYNPLVQSDNITRLMFAISFLTPLAIVCVVKIIRRTDKTEIKFAFLAVSLALALNG 106

Qy 61 VFTNTIKLIVGRPRDPFFYRCF 82

Db 107 VFTNTIKLIVGRPRDPFFYRCF 128

RESULT 5

US-10-103-313-452

Sequence 452, Application US/10103313

Publication No. US20030082758A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: P0207C1

CURRENT APPLICATION NUMBER: US/10/103,313

CURRENT FILING DATE: 2002-03-12

NUMBER OF SEQ ID NOS: 653

Prior Application removed - See File Wrapper or Palm

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 452

LENGTH: 149

TYPE: PRT

ORGANISM: Homo sapiens

US-10-103-313-452

Query Match 34.0%; Score 329; DB 15; Length 149;

Best Local Similarity 75.6%; Pred. No. 1.2e-31;

Matches 59; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

Qy 94 TGDKDVNVEGRKSPFSGHSSFAFAGLAFLAFASFLYLAGKLCFTPOGRGKSWRFCAFLSPLLF 153

```
Db 7 TGDPLVSEGRKSPFSSHSSFAFSGLGFTTFYLAGKLHCFTESGRKSWRLCAAILPLYC 66
Qy 154 AAVIALSRCTDYKHHWD 171
Db 67 AMIALSRMCDYKHHWD 84

RESULT 6
US-09-866-050A-170
; Sequence 170, Application US/09866050A
; Publication No. US20030040471A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866.050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 170
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Rat
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(123)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-866-050A-170

Query Match 21.0%; Score 203.5; DB 11; Length 123;
Best Local Similarity 45.2%; Pred. No. 1.3e-16;
Matches 47; Conservative 8; Mismatches 20; Indels 29; Gaps 3;

Qy 1 MWLYRNPVRAEYPTGPMFVIAFLSPLSLFLAKFLKKADTRDSR-----QACLAASL 54
Db 34 LWLYRNPVRAEYPTGPMFVIAFLSPLSLFLAKFLKKADTRDSR-----QACLAASL 54
Qy 55 ALAL-----NGVFTNTIKLVGRPRDPFFYRCFPDGLAHS 89
Db 94 CLYQHHTDSXQGHQAIAST-----ECSPXGIAHS 123

RESULT 7
US-10-106-698-5750
; Sequence 5750, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106.698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 5750
; LENGTH: 220
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
```

```
; LOCATION: (65)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (96)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (117)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-5750

Query Match 14.3%; Score 138; DB 15; Length 220;
Best Local Similarity 26.7%; Pred. No. 2.3e-08;
Matches 35; Conservative 24; Mismatches 58; Indels 14; Gaps 4;

Qy 53 SLALALNGVFTNTIKLVGRPRDPFFYRCFP-----DGLAHSDL-----MCTGDKDVVNE 102
Db 27 AFGLEATDIEFNAGQVVTGHLTPVFLTVCKPNVTSADCKAHHQFINNGNICTGDRVIEK 86
Qy 103 GRKSPFSGHSSFAFAGLAFASFYLAGKLHCFTPQGRKSWRFCAPLSPILFAAVIALSRT 162
Db 87 ARRSFSPKXHALSIYSALYATMYITSTIK--TXSSRLAKPVLCLGTLCTAFLTLGLNVR 142
Qy 163 CDYKHHWQDLL 173
Db 143 SEYRNHCSDVI 153

RESULT 8
US-09-860-670-125
; Sequence 125, Application US/09860670
; Patent No. US20020165137A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA127P1
; CURRENT APPLICATION NUMBER: US/09/860.670
; CURRENT FILING DATE: 2001-05-21
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 125
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (154)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (155)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (169)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (175)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-860-670-125

Query Match 11.9%; Score 115; DB 10; Length 180;
Best Local Similarity 23.2%; Pred. No. 1e-05;
Matches 29; Conservative 24; Mismatches 48; Indels 24; Gaps 4;

Qy 64 NTIKLVGRPRDPFFYRCFPDGLA-----HSDLMCTGDKDVVNEGRKSPSGHSS 113
Db 2 NAGQVVTGNLAPHFLALCKENYALGCGQYQTFISGEACTGNPDLLIMRAKTFPSKEA 61
Qy 114 FAFAGLAFASFYLAGKLHCFTPQGRKSWR-----FCAPLSPILFAAVIALSRTCDYKHH 168
Db 62 LSVVAAMVLTMYITNTI-----KAGTRLAKPVLCLGLMCLAF--LTGLNRAVEYRNH 112
Qy 169 WQDLL 173
```

```

; NAME/KEY: misc feature
; LOCATION: (178)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-103-313-340

Query Match 11.9%; Score 115; DB 15; Length 183;
Best Local Similarity 23.2%; Pred. No. 1.1e-05;
Matches 29; Conservative 24; Mismatches 48; Indels 24; Gaps 4;

QY 64 NTKILVGRPRDPFFRCFPPDGLA-----HSDLMCTGDKDVNVEGRKSPSGHSS 113
Db 5 NAGQVVTGNLAPHFLALCKPNYALGCCOYTQFISGEAEACTGNPDLMRARKTTPSKEAA 64
QY 114 FAFAGLAFASFVLAGKLHCFTQGRGKSWR-----FCAFLSPLFAAVIALSRCTCDYKH 168
Db 65 LSVYAMYLTMVITNTI-----KAKGTRLAKPVLCLGLMCLAF--LTGLNRVAEYRNH 115
QY 169 WODLL 173
Db 116 WSDVI 120

RESULT 11
US-10-103-313-493
; Sequence 493, Application US/10103313
; Publication No. US20030082758A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PZ07C1
; CURRENT APPLICATION NUMBER: US/10/103,313
; CURRENT FILING DATE: 2002-03-12
; NUMBER OF SEQ ID NOS: 653
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 493
; LENGTH: 203
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (73)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc feature
; LOCATION: (186)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc feature
; LOCATION: (190)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc feature
; LOCATION: (196)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-103-313-493

Query Match 11.4%; Score 110; DB 15; Length 203;
Best Local Similarity 23.2%; Pred. No. 4.8e-05;
Matches 33; Conservative 18; Mismatches 53; Indels 38; Gaps 5;

QY 60 GVFTNTI-----KLIVGRPRDPFFRC-----FPDGLAHSIDLMTCTGDK 97
Db 29 GLFTTIFANAGQVVTGNPTPHFLSVCRPNVTALGCLPPSPDRPXPDRFVTDQACAGSP 88
QY 98 DVNVEGRKSPSGHSSFAFAGLAFASFVLAGKLHCFTQGRGKSWR-----CAFLSPL 151
Db 89 SLVAARRAFPCDKAALCAVAVTYTMYV-----TLVFRVKSRLVKSLCLALLCPA 141
QY 152 LFAAVIALSRCTCDYKHWDLL 173
Db 142 FLVGWV---RVAEYRNHWSVDL 160

RESULT 12
US-10-156-761-15017

```

```

; Sequence 15017, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 15017
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-15017

```

```

Query Match          9.6%; Score 93; DB 15; Length 525;
Best Local Similarity 30.3%; Pred. No. 0.019;
Matches 46; Conservative 16; Mismatches 38; Indels 52; Gaps 8;

QY 34 AKPLKKADTDSRCAACLAALALNGVFTNTI-KLVGRPRDPFFYRCPPDGLAHSDLM 92
DB 44 AAVLATADGTRARRAARGVGSALASLTVNTVAKWATRRPRP-----LL 88
QY 93 CTGDKDVNNEGR-----KSPFSGHSSFAFA---GLAFASFYLAGKLHCFTPQGRGKS 141
DB 89 -----DLVPQIRHLTRPHITTSFSGHSASAAATGVALES-----TGVG-- 129
QY 142 WRFCAPLSPILLFAAVIALSRDCTYKHHWQDILL 173
DB 130 ----ALVAPL--AAAVAFSRVYGVHPGDVL 155

```

```

RESULT 13
US-09-864-761-40341
; Sequence 40341, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aescica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669

```

```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annotmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 40341
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL161744.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 14
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 13
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 14
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 11
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 10
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 12
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 14
; OTHER INFORMATION: SWISSPROT HIT: Q10022, EVALUATE 1.00e-05
; OTHER INFORMATION: EST_HUMAN HIT: AW157233.1, EVALUATE 3.00e-43
US-09-864-761-40341

Query Match          9.0%; Score 87; DB 9; Length 84;
Best Local Similarity 25.0%; Pred. No. 0.009;
Matches 21; Conservative 16; Mismatches 33; Indels 14; Gaps 2;

QY 53 SLALALNGVFTNTIKLVGRPRDPFFYRCPPDGLA-----HSDLMCTGDKDVNNE 102
DB 3 TFGLFATDIEFNAGQVVTGNLAPHFLALCKPNVTALGCCQYQTFISGEACTGNPDLMR 62
QY 103 GRKSPFSGHSSFAFAGLAFASFYL 126
DB 63 ARKTFPSKEAALS-----VYAAMYL 82

RESULT 14
US-10-156-761-9418
; Sequence 9418, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109

```

```
; SEQ ID NO 9418
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9418

Query Match      8.5%; Score 82; DB 15; Length 497;
Best Local Similarity 29.8%; Pred. No. 0.37; 36; Indels 52; Gaps 7;
Matches 42; Conservative 11; Mismatches 36; Indels 52; Gaps 7;

QY 45 SROACLAASIALALANGVFTNTI-KLIIVGRPRPDPFFYRCFPDGLAHSMDLCTGDKDVVNEG 103
Db 72 ARRAAARGLASLTLASATINTLGRSVRRPRVL-----DVPVHV 111
QY 104 RK-----SPPSGHSSFAFA---GLAFASFYLAGKLHCFTQGRGKSWRFCAFLSPLL 152
Db 112 RRLKRQPTTTSPPSGHSSFAFAAFAFATGVALES-----RGLG-----AALAPL- 152
QY 153 FAAVIALSRCTDYKHHQDILL 173
Db 153 -AAVALSRVYTVGVHFPDVL 172
```

```
RESULT 15
US-09-925-299-860
; Sequence 860, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 860
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (174)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-299-860
```

```
Query Match      8.0%; Score 77.5; DB 9; Length 184;
Best Local Similarity 25.8%; Pred. No. 0.35;
Matches 31; Conservative 13; Mismatches 55; Indels 21; Gaps 4;

QY 56 LALNGVFTNTIKLIIVGRPRPDPFFYRCFPDGLAHSMDLCTGDKDVVNEGKSPSGHSSFA 115
Db 12 LALNEGWNWLIKNNVQIEPRP-----CGGPHTAVGT-KYGMPSHSHSQFM 53
QY 116 FAGLAFASFYLAGKLHCFTQGR--GKSWRFCAFLSPLLFAAVIALSRCTDYKHHQDILL 173
Db 54 WFFSVYSFLFLYLRMH-QTNNAFLDILLRHLVLSGLIAVFLVSYRVLLIYHTWSQVL 112
```

Search completed: August 14, 2003, 13:16:39  
Job time : 57 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 14, 2003, 13:03:37 ; Search time 39 Seconds  
(without alignments)  
443.855 Million cell updates/sec

Title: US-09-876-997-399

Perfect score: 968

Sequence: 1 MWLYRNPYVEAEYFPTKPMF.....RTCDYKHHMODLLKCTNTAK 180

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	372	38.4	290	2 E86284	F9L1.2 protein - A
2	337	34.8	302	2 E84421	probable phosphatidic acid
3	247	25.5	289	2 S70114	probable membrane phosphatidic acid
4	242	25.0	279	2 T40445	probable membrane phosphatidic acid
5	200.5	20.7	274	2 S69561	probable membrane phosphatidic acid
6	143	14.8	341	2 T16951	probable membrane phosphatidic acid
7	124	12.8	239	2 S64327	protein F1386.5 [i
8	118	12.2	318	2 JE0159	conserved hypotet
9	118	12.2	318	2 H89632	hypothetical prote
10	115.5	11.9	186	1 D69206	hypothetical prote
11	112	11.6	218	2 E86750	hypothetical prote
12	104	10.7	217	2 D97672	conserved hypotet
13	104	10.7	242	2 AB2876	hypothetical prote
14	102.5	10.6	395	2 T24576	hypothetical prote
15	97.5	10.1	292	2 AF3647	phosphatidylglycer
16	93.5	9.7	213	2 T05449	hypothetical prote
17	93.5	9.7	217	2 AG1277	hypothetical prote
18	92	9.5	325	2 S66668	hydrogen peroxide-
19	90.5	9.3	330	1 F64346	hypothetical prote
20	89	9.2	259	2 AB7623	PAP2 homolog prote
21	87	9.0	1294	2 T19982	hypothetical prote
22	85.5	8.8	217	2 AG1640	hypothetical prote
23	83	8.6	322	2 T45568	hypothetical prote
24	78.5	8.1	238	2 E70524	hypothetical prote
25	78.5	8.1	260	2 D75404	phosphatidylglycer
26	77.5	8.0	378	2 C70435	hypothetical prote
27	77	8.0	184	2 A71046	hypothetical prote
28	77	8.0	274	2 AB2509	probable phosphatid
29	77	8.0	298	2 T13684	NADH2 dehydrogenas

ALIGNMENTS

RESULT 1

E86284

F9L1.2 protein - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 02-Nov-2001

C;Accession: E86284

R;Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;

C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: AB6141; MUID:21016719; PMID:11130712

A;Accession: E86284

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-290 <STO>

A;Cross-references: GB:AE005172; NID:95103807; PIDN:AAD39637.1; GSPDB:GN00141

C;Genetics:

A;Map position: 1

Query Match 38.4%; Score 372; DB 2; Length 290;

Best Local Similarity 47.8%; Pred. No. 7.2e-31;

Matches 75; Conservative 28; Mismatches 46; Indels 8; Gaps 4;

QY 21 VIAFLPSLIFLAKFLKADTRDSQACLAASLALANGVFTNTIKLIIVGRPRDPFYR 80

Db 73 LIAVLVFAVICVYFIRN-DVYDLHAILGLFSLVITGVITDAIKDVGRRPDPFWR 131

QY 81 CFPDGLA--HS---DLMCTGDKDVNKGKSPFGSHSSFAFAGLAFASFYLAGKLHCFPT 135

Db 132 CFPDGGIGIFHNVTKNVLCTGAKDVVKEGKSPFGSHSSFAFAGLAFASFYLAGKLHCFPT 135

QY 136 QGKSKWRFCAPLSPLFAAVIALSRCTDYKHHQDL 172

Db 190 DQRGHVAKLCIVILPLLVAALVGVSRVDDYWHHQQDV 226

RESULT 2

E84421

probable phosphatidic acid phosphatase [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001

C;Accession: E84421

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.;

euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

hypothetical prote  
hypothetical prote  
gene DN10 protein  
probable permease  
conserved hypotet  
conserved hypotet  
hypothetical prote  
hypothetical prote  
probable inner mem  
hypothetical prote  
nitrate transport  
PAP2 family protei  
serine/threonine-s  
hypothetical prote  
hypothetical prote  
hypothetical prote

RESULT 4  
T40445  
-threonine:4:0 acid phosphate homolog - fission yeast (Schizosaccharomyces pombe)



Matches	49;	Conservative	18;	Mismatches	59;	Indels	36;	Gaps	7;
---------	-----	--------------	-----	------------	-----	--------	-----	------	----

  

QY	14	FPTKPMFVIAFLPSLIFLAKFLKADTRDSRQACLAAS--LALALNGVFTNTIKLIVG	71
DB	63	FGDKGGIIVAVL--LALVIVLLFFKQKEAT-----IWFAATVWLSALNTV----	111
QY	72	RPRPDDFFVRCPPDGLAHSDLMCTGDKDVNNGRKSFPFGSHSSAFAGLAFASFVLAGKLH	131
DB	112	RERPD-IHRL-----AAAFANEAGKSPFGSHSIFATIIFGSIFFTCLGKL-	154
QY	132	CFTPQGRCKSWRFCAFLSPLLFAAVIALSRCTCDYKHHWQDLL	173
DB	155	-----KVRSSKILLGILICILIALIMFSRIFVGVHYPSDTI	190

  

RESULT 12

D97672  
 hypothetical protein AGR\_C 4732 [imported] - Agrobacterium tumefaciens (strain C58, Cere  
 C:Species: Agrobacterium tumefaciens  
 C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 12-May-2003  
 C:Accession: D97672  
 A.; Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,  
 R.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;  
 Science 294, 2323-2328, 2001  
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum  
 A:Reference number: A97359; MUID:21608551; PMID:11743194  
 A:Accession: D97672  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-217 <KUR>  
 A:Cross-references: GB:AE007869; PIDN:AAK88333.1; PID:gl5157811; GSPDB:GN00169  
 C:Genetics:  
 A:Gene: AGR\_C 4732  
 A:Map position: circular chromosome  
 C:Superfamily: bacitracin transport permease

  

Query Match	10.7%;	Score 104;	DB 2;	Length 217;
Best Local Similarity	31.1%;	Pred. NO. 0.0033;		
Matches	37;	Conservative	15;	Mismatches 41; Indels 26; Gaps 6;

  

QY	54	LALALNGVFTNTIKLIVGRPRDPFVRCFPDGLAHSDLMCTGDKDVNNGRKSFPFGSHSS	113
DB	79	LTIGTASILVHTLKEFLIGRPELFLMG-----AYSLTPFTGDNLY-----ESFPFGHST	129
QY	114	FAPAGLAFASFVLAGKLHCFTPQGRGKSWRFCAFLSPLLFAAVIALSRCTCDYKHHWQDL	172
DB	130	--AAGAGTGVFAM-----LMRPRF--WAF-----LLLALVIGSVRIVGAHYPSDV	171

  

RESULT 13

AB2897  
 conserved hypothetical protein Atu2611 [imported] - Agrobacterium tumefaciens (strain C58f  
 C:Species: Agrobacterium tumefaciens  
 C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 12-May-2003  
 C:Accession: AB2897  
 R.; Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.  
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell  
 ; Karp, P.; Romero, P.; Zhang, S.  
 Science 234, 2317-2323, 2001  
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
 ster, E.W.  
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
 A:Reference number: AB2577; MUID:21608550; PMID:11743193  
 A:Accession: AB2897  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-242 <KUR>  
 A:Cross-references: GB:AE008688; PIDN:AAL43592.1; PID:gl7741109; GSPDB:GN00186  
 A:Experimental source: strain C58 (Dupont)  
 C:Genetics:  
 A:Gene: Atu2611  
 A:Map position: circular chromosome  
 C:Superfamily: bacitracin transport permease

Query Match 10.7%; Score 104; DB 2; Length 242;  
Best Local Similarity 31.1%; Pred. No. 0.0037;  
Matches 37; Conservative 15; Mismatches 41; Indels 26; Gaps 6;  
QY 54 LALALNGVFTNTIKLIVGRPRDPFFKCPDGLAHSMLCTGDKDVNNEGRKSPPSGHSS 113  
DB 104 LTIGTASILVHTLFLIGRARPPELFEMG-----AYSLTPTGDNLY-----ESFPPSGHST 154  
QY 114 FAFAGLAFASGYLAGKLHCFTTPOGRGKSWRFCAPLSPLFAAVIALSRTCDYKHHWQDL 172  
DB 155 --AAGAFGVFAM-----LMPRRF---WAF-----LLALVIGSRVIVGAHYPSDV 196

## RESULT 14

T24576  
hypothetical protein T06D8.3 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C;Accession: T24576  
R;Palmer, S.  
submitted to the EMBL Data Library, April 1995  
A;Reference number: Z19909  
A;Accession: T24576  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-396 <WIL>  
A;Cross-references: EMBL:Z49130; PIDN:CAA88966.1; GSPDB:GN00020; CESP:T06D8.3  
A;Experimental source: clone T06D8  
C;Genetics:  
A;Gene: CESP:T06D8.3  
A;Map position: 2  
A;Introns: 11/3; 59/2; 184/2; 205/2; 233/3; 292/3; 340/2

Query Match 10.6%; Score 102.5; DB 2; Length 396;  
Best Local Similarity 23.7%; Pred. No. 0.0089;  
Matches 50; Conservative 29; Mismatches 81; Indels 51; Gaps 9;  
QY 3 LYRNPVVEAEY--PPTKP-MFVIAFLSPLSLFLAK-----FLKKAD----- 41  
DB 78 LYKFNFPEDENVVSVPLLYTLAFTIPPLILGCEVMFWLFTKPRKIVVANGCECPVH 137  
QY 42 --TRDSRQACLAALALANGVFTNTIKLIVGRPRDPFFYRC-----FPDGLAHS----- 89  
DB 138 LFTRRRLFRFVIYLAGLLIVQIFVDITIKMTGYQRPVFLSCNVISITACTAPLEHSPSPS 197  
QY 90 -DLMCT-GDKDVNNEGRKSPPSGHSSFAFAGLAFASGYLAGKLHCFTTPOGRGKSWRFCAP 147  
DB 198 PHLACNRYGADRLRYAWLTFPSLHAVSVSSAACFASLYIYMINL-----RGAPL 247  
QY 148 LSPLLFAAVIAL-----SRTCDYKHHWQDL 172  
DB 248 LRPLLIFGFMGLCIVDSFGRINGYKNHWRDI 278

## RESULT 15

AF3647  
phosphatidylglycerophosphatase (EC 3.1.3.27) [imported] - Brucella melitensis (strain 16  
C;Species: Brucella melitensis  
C;Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 03-Jun-2002  
C;Accession: AF3647  
R;DeIvecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,  
.; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess  
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis  
A;Reference number: AD3252; PMID:11756688  
A;Accession: AF3647  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-292 <KUR>  
A;Cross-references: GB:AE008918; PIDN:AAU54345.1; PID:gl7985327; GSPDB:GN00191  
A;Experimental source: strain 16M  
C;Genetics:

A;Gene: BMEI11103  
A;Map position: 11  
C;Keywords: phosphoric monoester hydrolase  
Query Match 10.1%; Score 97.5; DB 2; Length 292;  
Best Local Similarity 26.1%; Pred. No. 0.021;  
Matches 41; Conservative 25; Mismatches 50; Indels 41; Gaps 10;  
QY 24 FLSPLSLIFL-AKFLKKADTRDSRQACLAASLALALNGVFTNTIKLIVGRPRP-----DF 77  
DB 37 FLVPAALFVYFRFINRSPKSFN--CFFIIASIAVSIVIKILKIIFGRARPGVLDDG 94  
QY 78 FYCFDPDGLAHSMLCTGDKDVNNEGRKSPPSGHSSFAF-AGLAFASGYLAGKLHCFTTPO 136  
DB 95 FY-----GFTFFRL-----DREF-----NSFPSAHTGVAAGVALA---LIMQKH----- 132  
QY 137 GRGKSWRFCAFLSPLFAAVIALSRTCDYKHHWQDL 173  
DB 133 -----RWI-----VPIILGIVIASRIINAHYLSDDV 159

Search completed: August 14, 2003, 13:08:46  
Job time : 41 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 14, 2003, 13:03:36 ; Search time 22 Seconds  
(without alignments)  
384.764 Million cell updates/sec

Title: US-09-876-997-399

Perfect score: 968

Sequence: 1 MWLYRNPYVEAEYPTKPMF.....RTCDYKHHWQDLLKCTNTAK 180

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	172	17.8	379	1 WUN DROME	Q9V576 drosophila
2	143	14.8	341	1 YSX3 CAEEL	Q10022 caenorhabdi
3	124	12.8	239	1 YGIP YEAST	P53223 saccharomyc
4	90.5	9.3	330	1 Y374 METJA	Q57819 methanococc
5	77.5	8.0	378	1 YF69 AQUAE	Q67513 aquifex aeo
6	76.5	7.9	838	1 P HUMAN	Q04671 homo sapien
7	74.5	7.7	507	1 CENA EMENI	P22152 emericella
8	73	7.5	921	1 ITH4 PIG	P79283 sus scrofa
9	70.5	7.3	576	1 ML12 ARATH	O80961 arabidopsis
10	70	7.2	133	1 E1BS ADECT	P14265 canine aden
11	70	7.2	507	1 UDP2 PIG	P79303 sus scrofa
12	69	7.1	1704	1 ABC3 HUMAN	Q99758 homo sapien
13	69	7.1	2233	1 RRPL P13H4	P12577 human para
14	68	7.0	200	1 ISP2 BRAJA	P30961 bradyrhizob
15	68	7.0	276	1 YGFR ECOLI	Q46818 escherichia
16	67.5	7.0	237	1 YEIU ECOLI	P76445 escherichia
17	67.5	7.0	444	1 CRF1 HUMAN	P14998 homo sapien
18	67.5	7.0	551	1 FZD2 XENLA	Q39005 xenopus lae
19	67.5	7.0	1313	1 MIP1 SCHPO	P87141 schizosacch
20	67	6.9	305	1 ATNB TORCA	P05029 corpedo cal
21	67	6.9	826	1 R1N1 EBV	P03190 Epstein-bar
22	67	6.9	938	1 PM15 CHLPN	Q92883 chlamydia p
23	66.5	6.9	153	1 V66 HPV57	P22158 human papil
24	66.5	6.9	381	1 CYB CHLRE	P23662 chlamydomon
25	66.5	6.9	381	1 CYB CHLSM	P23663 chlamydomon
26	66.5	6.9	704	1 FNP BACSU	P50849 bacillus su
27	66.5	6.9	892	1 TNPA STRFR	P20189 streptomyce
28	66.5	6.9	2139	1 CCAC MOUSE	Q01815 mus musculu
29	66.5	6.9	2169	1 CCAC RAT	P22002 rattus norv
30	66.5	6.9	2221	1 CCAC HUMAN	Q13936 homo sapien
31	66	6.8	293	1 MT16 EMENI	P56859 emericella
32	66	6.8	344	1 GCP MYCTU	Q50709 mycobacteri
33	66	6.8	658	1 KPCI LYTP1	Q25378 lytechinus

## RESULT 1

ID	WUN DROME	STANDARD;	PRT;	379 AA.
AC	Q9V576; P91661; Q8MKU5; Q8WT60; Q9U9Y7;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Putative phosphatidate phosphatase (EC 3.1.3.4) (Phosphatidic acid phosphatase type 2) (Wunen protein) (Germ cell guidance factor).			
GN	WUN OR CG8804.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=7227;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM SHORT), FUNCTION, AND TISSUE SPECIFICITY.			
RX	MEDLINE=971138220; PubMed=8985246;			
RA	Zhang N., Zhang J.P., Purcell K.J., Chen Y., Howard K.;			
RT	"The Drosophila protein Wunen keeps migrating germ cells.";			
RL	Nature 385:64-67(1997).			
RL	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Berkeley;			
RC	MEDLINE=20196006; PubMed=10731132;			
RX	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,			
RA	George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,			
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,			
RA	Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,			
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,			
RA	Abriel J.P., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,			
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,			
RA	Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,			
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,			
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,			
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,			
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,			
RA	Dodson K., Dou L.B., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,			
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,			
RA	Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,			
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,			
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,			
RA	Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwack C.,			
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,			
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,			
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,			
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,			
RA	Merkulov G., Milshina N.V., Mobarri C., Morris J., Moshrefi A.,			
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,			
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacieb J.M.,			
RA	Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,			
RA	Reinart K., Remington K., Saunders R.D.C., Scheeler P., Shen H.,			
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,			
RA	Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,			
RA	Svirekas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,			

## ALIGNMENTS

34	66	6.8	1379	1 YFP9 SCHPO	Q14066 schizosacch
35	65.5	6.8	177	1 CCAC CHICK	O73707 gallus gall
36	65.5	6.8	380	1 METB HELPY	P56069 helicobacte
37	65.5	6.8	433	1 YBL4 YEAST	P38211 saccharomyc
38	65.5	6.8	1152	1 YBF7 CAEEL	Q21286 caenorhabdi
39	65.5	6.8	2171	1 CCAC RABIT	P15381 oryctolagus
40	64.5	6.7	247	1 LPXH_XANCP	P58976 xanthomonas
41	64.5	6.7	326	1 TRXB BORBU	P94284 borrelia bu
42	64.5	6.7	498	1 YM22 CAEEL	Q21339 caenorhabdi
43	64	6.6	255	1 THID AQUAE	O67772 aquifex aeo
44	64	6.6	368	1 PTIC ECOLI	P31452 escherichia
45	64	6.6	415	1 CRF1_MOUSE	P35347 mus musculu

Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,  
Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
"The genome sequence of *Drosophila melanogaster*";  
Science 287:2185-2195(2000).  
[3]  
REVIEWS, AND ALTERNATIVE SPLICING.  
STRAIN=Berkeley;  
MEDLINE=22426069; PubMed=12537572;  
Mitra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,  
Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,  
Harris N.B., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
Lewis S.E.;  
"Annotation of the *Drosophila melanogaster* euchromatic genome: a  
systematic review";  
Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
[4]  
SEQUENCE FROM N.A. (ISOFORM SHORT).  
STRAIN=Berkeley; TISSUE=Head;  
MEDLINE=20196012; PubMed=10731138;  
Rubin G.M., Hong L., Brokstein P., Evans-Holm M., Frise E.,  
Stapleton M., Harvey D.A.;  
"A *Drosophila* complementary DNA resource";  
Science 287:2222-2224(2000).  
[5]  
SEQUENCE FROM N.A. (ISOFORM LONG).  
STRAIN=Berkeley; TISSUE=Embryo;  
MEDLINE=22426066; PubMed=12537569;  
Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,  
George R.A., Guarin H., Kronmiller B., Pacleb J.M., Park S., Wan K.H.,  
Rubin G.M., Celniker S.E.;  
"A *Drosophila* full-length cDNA resource";  
Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).  
[6]  
SEQUENCE OF 22-99 FROM N.A.  
Russell C., Bartos S., Phillips R.G., Whittle R.;  
"Efficient functional dissection of enhancer sequences in *Drosophila*:  
a novel P- and hobo-based construct acting as an enhancer-capture  
element";  
Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
- FUNCTION: Responsible for guiding the germ cells early in the  
process of migration from the lumen of the developing gut towards  
the overlying mesoderm, where the germ cells enter the gonads. May  
be involved in lipid metabolism.  
- CATALYTIC ACTIVITY: A 3-sn-phosphatidate + H(2)O = a 1,2-diacyl-  
sn-glycerol + phosphate.  
- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
- ALTERNATIVE PRODUCTS:  
Event=Alternative splicing; Named isoforms=2;  
Name=Long;  
IsoId=Q9V576-1; Sequence=Displayed;  
Name=Short;  
IsoId=Q9V576-2; Sequence=VSP\_005084;  
Note=No experimental confirmation available;  
- TISSUE SPECIFICITY: Expressed in embryonic gut in a pattern that  
guides germ cells towards mesoderm (initially in hindgut and then  
on lower side of gut). During extended germ band stage, expressed  
in ectoderm as a medial band throughout the trunk.  
- SIMILARITY: BELONGS TO THE PHOSPHATIDATE PHOSPHATASE FAMILY.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL; U73822; AAC47449.1; -;  
DR EMBL; AE003833; AAF58942.2; -;  
DR EMBL; AE003833; AAF71066.1; -;  
DR EMBL; AF145595; AAD38570.1; -;  
DR EMBL; BT001729; AAN71484.1; -;  
DR EMBL; AY046533; AAL34392.1; -;  
DR FlyBase; FBgn0016078; wun.  
DR GO; GO:0016021; C:integral to membrane; NAS.  
DR GO; GO:0008195; F:phosphatidate phosphatase activity; NAS.  
DR GO; GO:0016311; P:dephosphorylation; NAS.  
DR GO; GO:0008354; P:germ-cell migration; IE.  
DR InterPro; IPR000326; PA\_FTPase.  
DR Pfam; PF01569; PAP2; 1.  
DR SMART; SM00014; acidppc; 1.  
KW Hydrolase; Developmental protein; Glycoprotein; Transmembrane;  
KW Alternative splicing.  
FT TRANSMEM 90 110 POTENTIAL.  
FT TRANSMEM 138 158 POTENTIAL.  
FT TRANSMEM 266 286 POTENTIAL.  
FT TRANSMEM 330 350 POTENTIAL.  
FT CARBOHYD 51 51 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 169 169 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT VARSPLIC 1 79 Missing (in isoform Short).  
FT CONFLICT 152 152 F -> L (IN REF. 4).  
FT CONFLICT 360 360 A -> G (IN REF. 4).  
SQ SEQUENCE 379 AA; 42789 MW; AD3C4F06B3DD9ADA CRC64;  
Query Match 17.8%; Score 172; DB 1; Length 379;  
Best Local Similarity 34.1%; Pred. No. 1.1e-10; Indels 32; Gaps 5;  
Matches 47; Conservative 14; Mismatches 45;  
QY 58 LNVGVTNTIKLVGRPRDPFFVRCFP---DGLAHSDLM-----CTG---DKDVVN 101  
DB 202 LSLQITDIAYKISGRLPHPFIACVQPMADGSTCDDAINAGKYIQFTCKGVSSARMK 261  
QY 102 EGRKSPFSGHSSFAFAGLAPASPYLAGKLHCFTTPQGRGKSWRCAFLSPLL-----FAA 155  
DB 262 EMRLSPFSGHSSFTFFAVVLYLYQARM-----TWRGSKLRHLQLFLFMVAV 311  
QY 156 VIALSRCTDYKHHWQDL 173  
DB 312 YTALSRVSDYKHHWSDVL 329  
RESULT 2  
YSX3 CAEEL  
ID YSX3 CAEEL STANDARD; PRT; 341 AA.  
AC Q10022;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE Hypothetical 39.0 kDa protein T28D9.3 in chromosome II.  
GN T28D9.3.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Waterston R.;  
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
CC -! SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
CC -! SIMILARITY: WEAK, TO YEAST D9719.9.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----







RT albinism (OCA2)";  
 RL Hum. Mutat. 12:434-434 (1998).  
 CC -!- FUNCTION: MAY BE INVOLVED IN THE TRANSPORT OF TYROSINE.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- DISEASE: DEFECTS IN OCA2 ARE THE CAUSE OF TYROSINASE-POSITIVE  
 CC OCULOCUTANEOUS ALBINISM (OCA-II OR OCA2). OCA2 IS AN AUTOSOMAL  
 CC RECESSIVE DISORDER OF PIGMENTATION IN THE SKIN, HAIR, AND EYES.  
 CC THE PHENOTYPE OF PATIENTS WITH OCA2 IS TYPICALLY SOMEWHAT LESS  
 CC SEVERE THAN IN THOSE WITH TYROSINASE-DEFICIENT OCA1. THERE ARE  
 CC SEVERAL FORMS OF OCA2, FROM TYPICAL OCA TO RELATIVELY MILD  
 CC "AUTOSOMAL RECESSIVE OCULAR ALBINISM" (AROA). OCA2 IS THE MOST  
 CC PREVALENT TYPE OF ALBINISM THROUGHOUT THE WORLD.  
 CC -!- SIMILARITY: BELONGS TO THE SLC13A FAMILY OF TRANSPORTERS.  
 CC P SUBFAMILY.  
 CC -!- DATABASE: NAME-Albinism database (ADB);  
 CC NOTE-OCA-II mutations;  
 CC WWW="http://www.cbc.umn.edu/tad/".  
 CC -!- DATABASE: NAME-Mutations of the P gene;  
 CC NOTE-Retina International's Scientific Newsletter;  
 CC WWW="http://www.retina-international.com/sci-news/psgenemut.htm".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; M99564; AAC36477.1; -;  
 DR EMBL; U19176; AAC13784.1; -;  
 DR EMBL; U19153; AAC13784.1; JOINED.  
 DR EMBL; U19154; AAC13784.1; JOINED.  
 DR EMBL; U19155; AAC13784.1; JOINED.  
 DR EMBL; U19157; AAC13784.1; JOINED.  
 DR EMBL; U19158; AAC13784.1; JOINED.  
 DR EMBL; U19159; AAC13784.1; JOINED.  
 DR EMBL; U19160; AAC13784.1; JOINED.  
 DR EMBL; U19161; AAC13784.1; JOINED.  
 DR EMBL; U19162; AAC13784.1; JOINED.  
 DR EMBL; U19163; AAC13784.1; JOINED.  
 DR EMBL; U19164; AAC13784.1; JOINED.  
 DR EMBL; U19165; AAC13784.1; JOINED.  
 DR EMBL; U19166; AAC13784.1; JOINED.  
 DR EMBL; U19167; AAC13784.1; JOINED.  
 DR EMBL; U19168; AAC13784.1; JOINED.  
 DR EMBL; U19169; AAC13784.1; JOINED.  
 DR EMBL; U19171; AAC13784.1; JOINED.  
 DR EMBL; U19172; AAC13784.1; JOINED.  
 DR EMBL; U19173; AAC13784.1; JOINED.  
 DR EMBL; U19174; AAC13784.1; JOINED.  
 DR EMBL; U19175; AAC13784.1; JOINED.  
 DR PIR; S28911; S28911.  
 DR Genew; HGNC:8101; OCA2.  
 DR MIM; 203200; -;  
 DR MIM; 203310; -;  
 DR GO; GO:0005737; C:cytoplasm; TAS.  
 DR GO; GO:0016021; C:integral to membrane; TAS.  
 DR GO; GO:0005215; F:transporter activity; TAS.  
 DR GO; GO:0005302; F:tyrosine transporter activity; TAS.  
 DR GO; GO:0006726; F:eye pigment biosynthesis; TAS.  
 DR InterPro; IPR000802; Afs\_pump.  
 DR TIGRFAMs; TIGR00935; 2a45; 1.  
 KW Transport; Transmembrane; Albinism; Disease mutation; Polymorphism.  
 FT DOMAIN 1 179 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 180 197 POTENTIAL.  
 FT DOMAIN 198 330 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 331 347 POTENTIAL.  
 FT DOMAIN 348 353 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 354 370 POTENTIAL.  
 FT DOMAIN 371 384 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 385 401 POTENTIAL.  
 FT DOMAIN 402 423 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 424 440 POTENTIAL.  
 FT DOMAIN 441 513 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 514 530 POTENTIAL.  
 FT DOMAIN 531 620 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 621 637 POTENTIAL.  
 FT DOMAIN 638 647 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 648 664 POTENTIAL.  
 FT DOMAIN 665 679 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 680 696 POTENTIAL.  
 FT DOMAIN 697 720 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 721 737 POTENTIAL.  
 FT DOMAIN 738 760 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 761 777 POTENTIAL.  
 FT DOMAIN 778 817 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 818 834 POTENTIAL.  
 FT DOMAIN 835 838 CYTOPLASMIC (POTENTIAL).  
 FT CARBOHYD 214 214 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 218 218 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 273 273 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 781 781 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARIANT 27 27 G -> R (IN OCA-II).  
 FT VARIANT 86 86 S -> R (IN OCA-II).  
 FT VARIANT 112 112 /FTid=VAR\_006118.  
 FT VARIANT 206 211 /FTid=VAR\_006119.  
 FT VARIANT 257 257 MISSING (IN OCA-II; SEVERE).  
 FT VARIANT 273 274 /FTid=VAR\_006120.  
 FT VARIANT 305 305 D -> A.  
 FT VARIANT 368 368 /FTid=VAR\_006121.  
 FT VARIANT 385 385 NW -> KV (IN OCA-II).  
 FT VARIANT 395 395 /FTid=VAR\_006122.  
 FT VARIANT 404 404 R -> W (IN OCA-II).  
 FT VARIANT 419 419 /FTid=VAR\_006123.  
 FT VARIANT 419 419 A -> V (IN OCA-II).  
 FT VARIANT 425 425 /FTid=VAR\_006124.  
 FT VARIANT 440 440 F -> I (IN OCA-II; SEVERE).  
 FT VARIANT 440 440 /FTid=VAR\_006125.  
 FT VARIANT 443 443 M -> L (IN OCA-II; SEVERE).  
 FT VARIANT 446 446 /FTid=VAR\_006126.  
 FT VARIANT 473 473 T -> M (IN OCA-II).  
 FT VARIANT 481 481 /FTid=VAR\_006127.  
 FT VARIANT 489 489 R -> Q.  
 FT VARIANT 549 549 /FTid=VAR\_007939.  
 FT VARIANT 443 443 L -> H.  
 FT VARIANT 446 446 V -> I (IN OCA-II).  
 FT VARIANT 473 473 /FTid=VAR\_006132.  
 FT VARIANT 481 481 M -> V (IN MILD OCA-II; AROA FORM).  
 FT VARIANT 489 489 I -> S (IN OCA-II).  
 FT VARIANT 481 481 /FTid=VAR\_006133.  
 FT VARIANT 489 489 A -> T (IN OCA-II).  
 FT VARIANT 549 549 /FTid=VAR\_007940.  
 FT VARIANT 549 549 N -> D (IN MILD/OCA-II; SEVERE).  
 FT VARIANT 549 549 /FTid=VAR\_006135.  
 FT VARIANT 549 549 H -> Q (IN OCA-II).  
 FT VARIANT 549 549 /FTid=VAR\_006136.

Query Match 7.9%; Score 76.5; DB 1; Length 838;  
 Best Local Similarity 25.2%; Pred. No. 3.7;  
 Matches 36; Conservative 17; Mismatches 57; Indels 33; Gaps 5;

Oy 43 RDSRQCLAAALALNGVETNTIKLVGRPRDPFFVRCPPDGLAHDLMCTGDKDVNE 102

Db 6 RDGRRYPGAPAVELLQTSVPSGLAELVAGRR-----LPRGAGGAD----- 46

QY 103 GRKSPF---SCHSFAFAGLAFASFYLAGLKHCFTPQGRGSKWRFCAFL--SPLLFAAVI 157  
 Db 47 PSHSCPAGAACQSWAPAGQEFASFLTKGRSHSLPQWSSSRKSDSCFTENTPLRLNSLQ 106  
 QY 158 ALSRTCDYKGH-----WQD 171  
 Db 107 EKGSRCPVYVHPFITAESWED 129

## RESULT 7

CRNA EMENI  
 ID CRNA EMENI STANDARD; PRT; 507 AA.  
 AC P22152;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DE 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Nitrate transporter (Nitrate permease).  
 GN CRNA.  
 OS Emericella nidulans (Aspergillus nidulans).  
 OC Eukaryota; Fungi; Ascomycota; Perizomycotina; Eurotiomycetes;  
 OC Eurotiales; Trichocomaceae; Emericella.  
 OX NCBI\_TaxID=162425;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91095428; PubMed=1986367;  
 RA Unkles S.E., Hawker K.L., Grieve C., Campbell E.I., Montague P.,  
 RA Kinghorn J.R.;  
 RT "crna encodes a nitrate transporter in Aspergillus nidulans.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 88:204-208(1991).  
 RN [2]  
 RP ERRATUM.  
 RA Unkles S.E., Hawker K.L., Grieve C., Campbell E.I., Montague P.,  
 RA Kinghorn J.R.;  
 RL Proc. Natl. Acad. Sci. U.S.A. 88:4564-4564(1991).  
 CC -|- FUNCTION: PERMEASE FOR NITRATE UPTAKE.  
 CC -|- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -|- INDUCTION: SUBJECT TO NITRATE AND NITRITE INDUCTION, AND NITROGEN  
 CC METABOLITE REPRESSION. CRNA EXPRESSION IS MEDIATED BY THE PRODUCTS  
 CC OF NIRA, AREA, AND NIAD.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; M61125; AAA62125.1; -;  
 DR EMBL; U34382; AAA76713.1; -;  
 DR InterPro; IPR007114; MPS.  
 DR InterPro; IPR004737; N03\_transporter.  
 DR InterPro; IPR005828; Sub\_transporter.  
 DR Pfam; PF00083; sugar tr; 1.  
 DR TIGRPFAMS; TIGR00886; 2A0108; 1.  
 KW Nitrate assimilation; Transport; Transmembrane.  
 FT DOMAIN 1 34 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 35 55 POTENTIAL.  
 FT DOMAIN 56 71 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 72 92 POTENTIAL.  
 FT DOMAIN 93 100 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 101 121 POTENTIAL.  
 FT DOMAIN 122 130 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 131 151 POTENTIAL.  
 FT DOMAIN 152 161 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 162 182 POTENTIAL.  
 FT DOMAIN 183 198 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 199 219 POTENTIAL.  
 FT DOMAIN 220 306 CYTOPLASMIC (HYDROPHILIC) (POTENTIAL).  
 FT TRANSMEM 307 327 POTENTIAL.  
 FT DOMAIN 328 357 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 358 378 POTENTIAL.

FT DOMAIN 379 389 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 390 410 POTENTIAL.  
 FT DOMAIN 411 417 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 418 438 POTENTIAL.  
 FT DOMAIN 439 507 CYTOPLASMIC (POTENTIAL).  
 SQ SEQUENCE 507 AA; 54925 MW; 4A3D3FA643F16952 CRC64;  
 Query Match 7.7%; Score 74.5; DB 1; Length 507;  
 Best Local Similarity 22.8%; Pred. No. 3.4; Indels 71; Gaps 12;  
 Matches 49; Conservative 31; Mismatches 64;  
 QY 20 FVIAFLS-----PLSLIFLAKFLKKADTRDS-----RQAC----- 49  
 Db 43 FMLAFLSWYAFPPLLTVTIRDDLSQSQIANSNIALLATLLVRLICGPGCDRFGRLV 102  
 QY 50 -----LAASLALALNGVNTIKLIVGRPRDPFRCFPGDLAHSIDLMCTG--DKQVYNE 102  
 Db 103 FIGILLVGSITPAMAGLVTSPQGLIALR-----FFIGILGTFVPCQVWCTGFDFKSIV-- 156  
 QY 103 GRKSPFGHSSFAFAGLAF-----SFYLAKLKHCFTPQGRG-----KSWRFCAFLSP--LL 152  
 Db 157 -----GTANSLAAGLGNAGGIVFVMPAIFDSLRDQGLPAHKAWR-VAVIVPILI 208  
 QY 153 FAAVIALSRITCD-----YKHHWQDLKCTNTA 179  
 Db 209 VAAALGMLFTCDDTPTGKWSERHIW--MKEDTQTA 241

## RESULT 8

ITH4 PIG  
 ID ITH4 PIG STANDARD; PRT; 921 AA.  
 AC P79263;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Inter-alpha-trypsin inhibitor heavy chain H4 precursor (ITI heavy  
 DE chain H4) (Inter-alpha-inhibitor heavy chain 4) (Inter-alpha-trypsin  
 DE inhibitor family heavy chain-related protein) (IHRP) (Major acute  
 DE phase protein) (MAP).  
 GN ITH4 OR IHRP.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 28-36; 695-703 AND 710-722.  
 RC TISSUE=Liver;  
 RX MEDLINE=96271024; PubMed=8830057;  
 RA Hashimoto K., Tobe T., Sumiya J.-I., Sano Y., Choi-Miura N.-H.,  
 RA Ozawa A., Yasue H., Tomita M.;  
 RT "Primary structure of the pig homologue of human IHRP: inter-alpha-  
 RT trypsin inhibitor family heavy chain-related protein.";  
 RL J. Biochem. 119:577-584(1996).  
 RN [2]  
 RP PRELIMINARY SEQUENCE OF 267-556 FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=90371455; PubMed=1697703;  
 RA Buchman T.G., Cabin D.E., Vickers S., Deutschman C.S., Delgado E.,  
 RA Sussman M.M., Bulkley G.B.;  
 RT "Molecular biology of circulatory shock. Part II. Expression of four  
 RT groups of hepatic genes is enhanced after resuscitation from  
 RT cardiogenic shock.";  
 RL Surgery 108:559-566(1990).  
 RN [3]  
 RP SEQUENCE OF 28-54 AND 223-240.  
 RC TISSUE=Serum;  
 RX MEDLINE=96013138; PubMed=7556597;  
 RA Gonzalez-Ramon N., Alava M.A., Sarsa J.A., Pineiro M., Escartin A.,  
 RA Garcia-Gil A., Lampreave F., Pineiro A.;  
 RT "The major acute phase serum protein in pigs is homologous to human  
 RT plasma kallikrein sensitive PK-120.";  
 RL FEBS Lett. 371:227-230(1995).  
 CC -|- FUNCTION: MAY BE INVOLVED IN ACUTE PHASE REACTIONS.

CC -1- TISSUE SPECIFICITY: Liver-specific.  
 CC -1- INDUCTION: LEVELS INCREASE SIGNIFICANTLY AFTER CARDIOGENIC  
 CC SHOCK.  
 CC -1- PTM: APPEARS TO BE BOTH N- AND O-GLYCOSYLATED (BY SIMILARITY).  
 CC -1- PTM: CLEAVED BY PLASMA KALLIKREIN TO YIELD 55- AND 25-KDA  
 CC FRAGMENTS.  
 CC -1- SIMILARITY: BELONGS TO THE ITIH FAMILY.  
 CC -1- SIMILARITY: Contains 1 VWFA domain.  
 CC -1- CAUTION: REF.2 SEQUENCE IS INCORRECT DUE TO FRAMESHIFTS AND OTHER  
 CC SEQUENCING ERRORS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U43164; AAD00024.1; -;  
 CC EMBL; S82800; AAB46821.1; -;  
 CC EMBL; M29507; -; NOT ANNOTATED\_CDS.  
 CC PIR; JC4625; JC4625;  
 CC InterPro; IPR006587; VIT.  
 CC InterPro; IPR002035; VWF\_A.  
 CC Pfam; PF00092; vwa; 1.  
 CC SMART; SM00609; VIT; 1.  
 CC SMART; SM00327; VWF; 1.  
 CC PROSITE; PS50234; VWFA; 1.  
 CC Serine protease inhibitor; Repeat; Signal; Multigene family;  
 KW Glycoprotein.  
 KW SIGNAL 1 27  
 FT CHAIN 28 921  
 FT FT  
 FT FT  
 FT DOMAIN 270 428  
 FT CARBOHYD 80 80  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 205 205  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 242 242  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 513 513  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 577 577  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 49 50  
 FT HT -> SK (IN REF. 3).  
 FT CONFLICT 703 703  
 FT D -> H (IN REF. 1; AA SEQUENCE).  
 FT CONFLICT 921 AA; 102146 MW; E2BF95925DE8D07C CRC64;  
 SQ SEQUENCE 921 AA; 102146 MW; E2BF95925DE8D07C CRC64;  
 Query Match 7.5%; Score 73; DB 1; Length 921;  
 Best Local Similarity 25.3%; Pred. No. 9.7;  
 Matches 40; Conservative 30; Mismatches 56; Indels 32; Gaps 8;  
 QY 10 EAEVFTKMPF-----VIAFLSPLSLIFLAKFLKADTRDSROACLAASLALALNGVF- 62  
 DB 526 EAEFLSPKIFHSPMERLWAYLTITQQL--LAQTVASDA--EKKALPARALSLSLNSYFV 581  
 QY 63 TMTIKLVGRPRDPFFYRCPPDGLAHLDMCTGDKDVNVEGRK-SPPSGHSSFAFAGLAF 121  
 DB 582 TPLTSMWITK-----PEGOEQSQV---AEKPVENGRQGNTHSGHSSQFHSVGD 628  
 QY 122 ASFYLAGK-----LHCFTPOGRGKSWRFCAFLSPLL 152  
 DB 629 RTSRLTGGSSVDVPVFSHRRGWKMGQAQGFEXKSYLPPEL 666  
 RESULT 9  
 ML12 ARATH  
 ID ML12 ARATH STANDARD; PRT; 576 AA.  
 AC O80961; Q94KB3;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE MLO-like protein 12 (ACM1012).  
 GN MLO12 OR ATG39200 OR T16B24.16.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Devoto A., Hartmann H.A., Piffanelli P., Elliott C., Simmons C.,  
 RA Taramino G., Goh C., Cohen F.E., Schulze-Liefert P., Panstruga R.;  
 RA "Molecular phylogeny and domain-specific co-evolution of the  
 RT plant-specific seven transmembrane MLO family.";  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=20083487; Pubmed=10617197;  
 RA Lin X., Kaul S., Roundley S.D., Shea T.P., Benito M.-I., Town C.D.,  
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,  
 RA Buell C.R., Kronum K.A., Lee J.J., Ronning C.M., Koo H.L.,  
 RA Moffat K.S., Cronin L.A., Shen M., Pai G., Van Aken S., Umayam L.,  
 RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,  
 RA Goodman H.M., Somerville C.R., Copenhagen G.P., Preuss D.,  
 RA Niernman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,  
 RA Venter J.C.;  
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis  
 RL thaliana.";  
 RL Nature 402:761-768 (1999).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE MLO FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF369573; AAK53805.1; -;  
 CC EMBL; AC004697; AAC28997.2; -;  
 CC InterPro; IPR004326; MLO.  
 CC Pfam; PF03094; MLO; 1.  
 CC Transmembrane; Multigene family.  
 FT TRANSMEM 14 34  
 FT POTENTIAL.  
 FT TRANSMEM 60 80  
 FT POTENTIAL.  
 FT TRANSMEM 156 176  
 FT POTENTIAL.  
 FT TRANSMEM 279 299  
 FT POTENTIAL.  
 FT TRANSMEM 309 329  
 FT POTENTIAL.  
 FT TRANSMEM 363 383  
 FT POTENTIAL.  
 FT TRANSMEM 406 426  
 FT POTENTIAL.  
 SQ SEQUENCE 576 AA; 66548 MW; 43DA9F6AED64D8E6 CRC64;  
 Query Match 7.3%; Score 70.5; DB 1; Length 576;  
 Best Local Similarity 22.2%; Pred. No. 10;  
 Matches 38; Conservative 23; Mismatches 57; Indels 53; Gaps 8;  
 QY 20 FVIAFLSPL---SLIFLAKFLKADTRDSROACLAASLALALNGVFTNTIKLVGRPRPD 76  
 DB 21 FVLLFTSIMIEYFLHFGHFWFKKKKALKSEALKKAEKLMILG-FISLLVLQTVSE 79  
 QY 77 FFYRCPPDGLA-----HSDLMCTGDKDVNVEGRK-----SPPSGHSSFAFAG-- 118  
 DB 80 I---CIPRNIAATWHPCSNHOEIAKYG-KDYIDGRKILEDSDNSDFSPRNLAATGYD 135  
 QY 119 -----LAFASFYLAGKLHCF-----TPQGRGKSW 142  
 DB 136 KCAEKGKVALVSAYGIHQHIFVLAVFHVLYCIITYALGKTKMKKWSW 186  
 RESULT 10  
 E1BS ADECT  
 ID E1BS ADECT STANDARD; PRT; 133 AA.  
 AC P14265;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE E1B protein, small T-antigen (Early E1B 15 kDa protein).  
OS Canine adenovirus type 2 (strain Toronto A 26-61).  
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.  
OX NCBI\_TaxID=69152;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90021176; PubMed=2800332;  
RA Shibata R., Shinagawa M., Iida Y., Tsukiyama T.;  
RT "Nucleotide sequence of E1 region of canine adenovirus type 2.";  
RL Virology 172:460-467(1989).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL: J04368; AAA42471.1; --  
CC F1R; B34165; WMADN2.  
CC InterPro: IPR002924; AdenoE1B\_19kDa.  
CC InterPro: IPR002475; BCL2 family.  
CC Pfam: PF01691; Adeno E1B 19K; 1.  
CC ProDom: PD004074; AdenoE1B\_19kDa; 1.  
CC PROSITE: PS50062; BCL2\_FAMILY; 1.  
CC Early protein.  
KW Early protein.  
SQ SEQUENCE 133 AA; 15219 MW; 2C733F250830B504 CRC64;  
Query Match 7.2%; Score 70; DB 1; Length 133;  
Best Local Similarity 35.7%; Pred. No. 2.2;  
Matches 30; Conservative 9; Mismatches 31; Indels 14; Gaps 6;  
QY 75 PDFYR-CFPGDGLSHDLMTGCDKDVNVEGR--KSPFSGHSSFAFAGLAFASFLVLAGKLH 131  
DB 24 PGFRRFCFP---ALADVV--GNIVEGEGRFWOLPENH---AFWGLLRGFTVAGFTE 75  
QY 132 CFTP---QGRGKSWRFCAFLSPLL 152  
DB 76 IITAAQLNQRQLAFIAFLSPLL 99  
RESULT 11  
UDP2\_PIG STANDARD; PRT; 507 AA.  
AC P79303;  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE UTP-glucose-1-phosphate uridylyltransferase 2 (EC 2.7.7.9) (UDP-  
DE glucose pyrophosphorylase 2) (UDPGP 2) (UGPase 2).  
GN UGP2.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Muscle;  
RA Looft C., Paul S.;  
RT "cDNA sequencing of the porcine UDP glucose pyrophosphorylase  
RT gene.";  
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: PLAYS A CENTRAL ROLE AS A GLUCOSYL DONOR IN CELLULAR  
CC METABOLIC PATHWAYS. UTP + alpha-D-glucose 1-phosphate =  
CC diphosphate + UDP-glucose.  
CC -1- SUBUNIT: Homooctamer (By similarity).  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -1- SIMILARITY: BELONGS TO THE EUKARYOTIC UDGP FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL: X99312; CAA67690.1; --  
CC InterPro: IPR002618; UDGP.  
CC Pfam: PF01704; UDGP; 1.  
KW Transferase; Kinase; Nucleotidyltransferase; Multigene family.  
FT INIT MET 0 0 BY SIMILARITY  
SQ SEQUENCE 507 AA; 56821 MW; 327F66D104D53965 CRC64;  
Query Match 7.2%; Score 70; DB 1; Length 507;  
Best Local Similarity 27.6%; Pred. No. 10;  
Matches 27; Conservative 14; Mismatches 27; Indels 30; Gaps 6;  
QY 97 KDVNVEGRKS---FPGHSSFAFAGLAFASFLVLAGKLHCFPGKSWRFCAFLSPLL 153  
DB 205 KDVSYSGENTEAWYPPGH-----GDIYASFYNSGLDFTI--GEGKEYIFVSNIDNL-- 254  
QY 154 AAVIAL-----SRTCDYKHHWQDLLKCTNTAK 180  
DB 255 GATVDLYILNHLNMPNPGRCPEF-----VMEATNKAR 286  
RESULT 12  
ABC3\_HUMAN STANDARD; PRT; 1704 AA.  
AC Q99758; Q92473;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE ATP-binding cassette, sub-family A, member 3 (ATP-binding cassette  
DE transporter 3) (ATP-binding cassette 3) (ABC-C transporter).  
GN ABC3 OR ABC3.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Thyroid carcinoma;  
RX MEDLINE=96326608; PubMed=8706931;  
RA Klugbauer N., Hofmann F.;  
RT "Primary structure of a novel ABC transporter with a chromosomal  
RT localization on the band encoding the multidrug resistance-associated  
RT protein.";  
RL FEBS Lett. 391:61-65(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97179225; PubMed=9027511;  
RX Connors T.D., van Raay T.J., Petry L.R., Klinger K.W., Landes G.M.,  
RA Burn T.C.;  
RA "The cloning of a human ABC gene (ABC3) mapping to chromosome  
RT 16p13.3.";  
RL Genomics 39:231-234(1997).  
CC -1- FUNCTION: MAY BE A TRANSPORTER, ITS NATURAL SUBSTRATE HAS NOT BEEN  
CC FOUND YET (BY SIMILARITY). MAY ACT AS AN EFFLUX PUMP FOR  
CC CHEMOTHERAPEUTICS DRUGS.  
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN LUNG, FOLLOWED BY BRAIN,  
CC PANCREAS, SKELETAL MUSCLE AND HEART. WEAKLY EXPRESSED IN PLACENTA,  
CC KIDNEY AND LIVER. ALSO EXPRESSED IN MEDULLARY THYROID CARCINOMA  
CC CELLS (MTC) AND IN C-CELL CARCINOMA.  
CC -1- DOMAIN: MULTIFUNCTIONAL POLYPEPTIDE WITH TWO HOMOLOGOUS HALVES,  
CC EACH CONTAINING AN HYDROPHOBIC MEMBRANE-ANCHORING DOMAIN AND AN  
CC ATP BINDING CASSETTE (ABC) DOMAIN (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. ABCA SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; U78735; AAC50967.1; --  
 EMBL; X97187; CA65825.1; --  
 PIR; A59188; A59188.  
 PIR; S71363; S71363.  
 Genew; HGNC:33; ABCA3.  
 MIM; 601615; --  
 GO; GO:0016021; C: integral to membrane; TAS.  
 GO; GO:0005624; C: membrane fraction; TAS.  
 GO; GO:0004009; F: ATP-binding cassette (ABC) transporter acti. . . ; TAS.  
 GO; GO:0005215; F: transporter activity; TAS.  
 GO; GO:0009315; P: drug resistance; TAS.  
 GO; GO:0006832; P: small molecule transport; TAS.  
 InterPro; IPR003593; AAA ATPase.  
 InterPro; IPR003439; ABC transporter.  
 Pfam; PF00005; ABC\_tran; 2.  
 ProDom; PD000006; ABC\_transporter; 2.  
 SMART; SM00382; AAA; 2.  
 PROSITE; PS00211; ABC\_TRANSPORTER\_1; 1.  
 PROSITE; PS00893; ABC\_TRANSPORTER\_2; 2.  
 ATP-binding; Transport; Transmembrane.  
 TRANSMEM 22 42 POTENTIAL.  
 TRANSMEM 249 269 POTENTIAL.  
 TRANSMEM 307 327 POTENTIAL.  
 TRANSMEM 344 364 POTENTIAL.  
 TRANSMEM 373 393 POTENTIAL.  
 TRANSMEM 405 425 POTENTIAL.  
 TRANSMEM 447 467 POTENTIAL.  
 TRANSMEM 925 945 POTENTIAL.  
 TRANSMEM 1100 1120 POTENTIAL.  
 TRANSMEM 1144 1164 POTENTIAL.  
 TRANSMEM 1183 1203 POTENTIAL.  
 TRANSMEM 1213 1233 POTENTIAL.  
 TRANSMEM 1245 1265 POTENTIAL.  
 TRANSMEM 1306 1326 POTENTIAL.  
 NP\_BIND 566 573 ATP (POTENTIAL).  
 NP\_BIND 1416 1423 ATP (POTENTIAL).  
 CONFLICT 36 36 P -> S (IN REF. 2).  
 CONFLICT 196 196 L -> P (IN REF. 2).  
 CONFLICT 196 196 L -> P (IN REF. 2).  
 SSQUENCE 1704 AA; 191387 MW; AF0098DAF7A04F5F CRC64;

Query Match 7.1%; Score 69; DB 1; Length 1704;  
 Best Local Similarity 26.1%; Pred. No. 52;  
 Matches 35; Conservative 11; Mismatches 40; Indels 48; Gaps 7;  
 QY 51 AASLALALNGVFTNTIKLVGRPRDPFFRCPPDGLAHSIDLMCTGDKDVVNEGRKSP--- 107  
 DB 1051 ATALAVVDNLLF-----KLLCG-PHASTVWSNFPQ--PRSAALQAA--KQDFNEGRKGPDA 1101  
 QY 108 -----PSGHSFSAFAGLAFASFYLAGKLHCFTPQGRGKSWR 143  
 DB 1102 LNLFAFAFLASTSILAVSERAVQAKHQVFV-SGVHVSFWLSALL-----WD 1149  
 QY 144 FCAFLSPLFAAVI 157  
 DB 1150 LISFLIPSLLLVV 1163

RESULT 13  
 RRPL PI3H4  
 ID RRPL PI3H4 STANDARD; PRT; 2233 AA.  
 AC PI2577;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE RNA polymerase beta subunit (EC 2.7.7.48) (Large structural protein)  
 DE (L protein).  
 GN L.  
 OS Human parainfluenza 3 virus (strain NIH 47885).  
 OS Viruses; ssRNA negative-strand viruses; Mononegavirales;

OC Paramyxoviridae; Paramyxovirinae; Respirovirus.  
 OX NCBI\_TaxID=11217;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88306242; PubMed=2841798;  
 RA Galinski M.S., Mink M.A., Pons M.W.;  
 RT "Molecular cloning and sequence analysis of the human parainfluenza 3  
 RT virus gene encoding the L protein.";  
 RL Virology 165:499-510(1988).  
 RN [2]  
 RP SEQUENCE OF 1-26 FROM N.A.  
 RX MEDLINE=88032139; PubMed=2822598;  
 RA Storey D.G., Cote M.-J., Dimock K., Kang C.Y.;  
 RT "Nucleotide sequence of the coding and flanking regions of the human  
 RT parainfluenza virus 3 hemagglutinin-neuraminidase gene: comparison  
 RT with other paramyxoviruses";  
 RL Intervirology 27:69-80(1987).  
 CC -!- FUNCTION: PROBABLE COMPONENT OF THE ACTIVE POLYMERASE. IT MAY  
 CC FUNCTION IN MRNA SYNTHESIS, CAPPING, METHYLATION AND POLY(A)  
 CC SYNTHESIS OF NEWLY SYNTHESIZED VIRAL MRNAs, RNA EDITING OF THE P  
 CC GENE TRANSCRIPT, AND PROTEIN KINASE ACTIVITY.  
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
 CC (RNA)(N).  
 CC -!- SIMILARITY: WITH L PROTEIN OF OTHER PARAMYXOVIRUSES.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; M21649; AAA46854.1; --  
 EMBL; M20402; AAA46857.1; --  
 PIR; B46451; B46451.  
 InterPro; IPR007098; RNA\_pol monom.  
 InterPro; IPR001016; Viral\_RNA\_pol\_L.  
 Pfam; PF00946; Paramyx\_RNA\_pol; 1.  
 KW Transferase; RNA-directed RNA polymerase.  
 FT CONFLICT 26 26 K -> R (IN REF. 2).  
 FT CONFLICT 26 26 K -> R (IN REF. 2).  
 SSQUENCE 2233 AA; 255802 MW; DEBBB8B5DE74B4638 CRC64;

Query Match 7.1%; Score 69; DB 1; Length 2233;  
 Best Local Similarity 32.7%; Pred. No. 71;  
 Matches 16; Conservative 10; Mismatches 23; Indels 0; Gaps 0;

QY 76 DFYRCPPDGLAHSIDLMCTGDKDVVNEGRKSPFSGHSFSAFAGLAFASF 124  
 DB 1576 DLFMRWLGVSLEIYICSDMEVANDRKOAFISRHLFSVCCLAETASF 1624

RESULT 14  
 ISPZ BRAJA  
 ID ISPZ BRAJA STANDARD; PRT; 200 AA.  
 AC P30961;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Probable intracellular septation protein.  
 GN ISPZ OR BL0472.  
 OS Bradyrhizobium japonicum.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Bradyrhizobiaceae; Bradyrhizobium.  
 OX NCBI\_TaxID=375;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=USDA 110;  
 RX MEDLINE=22484998; PubMed=12597275;  
 RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,  
 RA Sasamoto S., Watanabe A., Idesawa K., Iriuchi M., Kawashima K.,  
 RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,  
 RA Tabata S.;



